



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number 128108

To: Sarvamangala Devi

Location: REM 3C18

Art Unit: 1645

Wednesday, August 04, 2004

Case Serial Number: 09/700712

From: Beverly Shears

Location: Remsen Bldg.

RM 1A54

Phone: 571-272-2528

beverly.shears@uspto.gov

Search Notes

Shears, Beverly

128108

From: Devi, Sarvamangala
Sent: Monday, July 26, 2004 12:29 PM
To: Shears, Beverly
Subject: 09/700,712

Beverly:

Would you please perform a search for SEQ ID NO: 1 in commercial sequence and interference databases in application SN 09/700,712?

Please include a search for the inventors: Nils Carlin and Michael R. Lebens; and a text search for: A thy A or thymine A strain of Vibrio cholerae comprising a site directed mutagenesis in the thy A gene.

Thanks.

S. DEVI, Ph.D.
AU 1645
Rems - 3C18



STAFF USE ONLY

Date completed: 08-02-04

Searcher: Beverly 2528

Terminal time: 20 / 11

Elapsed time: _____

CPU time: _____

Total time: 23 / 21

Number of Searches: _____

Number of Databases: 3

Search Site

____ STIC

____ CM-1

____ Pre-S

Type of Search

____ N.A. Sequence

____ A.A. Sequence

____ Structure

____ Bibliographic

Vendors

____ IG

____ STN

____ Dialog

____ APS

____ Geninfo

____ SDC

____ DARC/Questel

____ Other CGW

This Page Blank (uspis,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 12:17:10 ; Search time 4825 Seconds

(without alignments)
18003.955 Million cell updates/sec

Title: US-09-700-712A-1

Perfect score: 2909 1 gagaaggttgcgtcgcgc.....cagcaaatgcgcgtgcag 2909

Sequence: 1 gagaaggttgcgtcgcgc.....cagcaaatgcgcgtgcag 2909

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estbm:*
3: em_estbn:*
4: em_estbu:*
5: em_estbv:*
6: em_estbl:*
7: em_estbr:*
8: em_estro:*
9: gb_estl:*
10: gb_est2:*
11: gb_estc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	249.8	8.6	1128	28	BZ577690 msh2_5528
2	194.8	6.7	805	28	BZ556310 pacsl-60
3	191.8	6.6	853	28	BZ578932 msh2_6070
4	161	5.5	1141	28	BZ559267 pacsl-164

Result No.	Score	Query Match	Length	DB ID	Description
5	152	5.2	1375	28	BZ553447 pacsl-60
6	115.6	4.0	406	28	BZ553227 pseu
7	111.2	3.8	1034	28	BZ561697 pacsl-164
8	111	3.8	878	14	CD379073 PTM04242
9	111	3.8	921	14	CD382942 PTM08112
10	106.2	3.7	855	14	CD376447 PTM01614
11	104	3.6	843	14	CA588286 hab49d11
12	103.8	3.6	813	14	CA588344 hab50c01
13	103.2	3.5	828	14	CA475083 AGENCOURT
14	96	3.3	685	14	BW039803 BW039803
15	96	3.3	686	13	BW039194 BW039194
16	96	3.3	801	14	CA588156 hab47903
17	92.4	3.2	662	14	CA589861 hab39c07
18	89	3.1	522	13	BW061136 BW061136
19	89	3.1	632	13	BW295886 BW295886
20	88.8	3.1	524	14	CD336403 STRFUS37
21	86.6	3.0	591	12	BG732886 346545 MA
22	85	2.9	683	10	AW423104 f168d06.X
23	84.8	2.9	700	14	CA353127 624588 NC
24	84.8	2.9	712	14	CA373071 647095 NC
25	84.6	2.9	483	29	CG637075 OGT362450
26	84.6	2.9	606	10	AW546106 L0003F03
27	84.6	2.9	668	13	BY746793 BY746793
28	84.6	2.9	976	13	BY746114 BY746114
29	84.6	2.9	2255	11	AK088606 Mus muscu
30	84.6	2.9	2256	11	AK088134 Mus muscu
31	84.6	2.9	3128	11	AK035663 Mus muscu
32	84	2.9	677	14	CB420463 593420 MA
33	83.8	2.9	756	12	B1078307 602872658
34	82.8	2.8	776	14	CF995901 AGENCOURT
35	82.4	2.8	528	14	CD330233 STRFUS37
36	82	2.8	748	13	B1055160 UI-M-FD0
37	81.8	2.8	673	14	CD313851 STRFUS37
38	81.4	2.8	407	14	CD330290 STRFUS37
39	81.4	2.8	1121	11	AK011435 Mus muscu
40	80.8	2.8	945	10	BF783860 602109784
41	80.8	2.8	2486	11	AK004832 Mus muscu
42	80.2	2.8	952	10	BF782519 602106535
43	80.2	2.8	1024	12	BM924499 AGENCOURT
44	79.8	2.7	974	10	BF607905 MY1.00084
45	79.2	2.7	567	9	AW108416 um3004.X

ALIGNMENTS

RESULT 1
BZ577690 1128 bp DNA linear GSS 17-DEC-2002
LOCUS msh2_5528.y2 msh Pseudomonas aeruginosa genomic clone msh2_5528,
DEFINITION genomic survey sequence.
ACCESSION BZ577690
VERSION BZ577690.1 GI:27212751
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1128)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,U.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence Variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) in press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers

source
1..1128
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="msb2_5528"
/clone.lib="msb"
/note="Environmental isolate. Whole genomic shotgun library."

ORIGIN

Query Match 8.6%; Score 249.8; DB 28; Length 1128;
Best Local Similarity 59.7%; Pred.No. 1.9e-61;
Matches 453; Conservative 0; Mismatches 301; Indels 5; Gaps 2;
29 CTGAGTTTCCCATATTGACCCCGTATGTTTTCATGAGCCCTCAGGCGCGCTG 88
93 CTGAGTATCCCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 152
89 TATGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 148
153 TACGGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 212
149 GATGCGCGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 208
213 AAGCGCTTGAACCGACCTGACCAAGAGGCGCTTCCAGACCTGCTTCTGGTGC 272
209 TTAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 268
273 TCGCGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 332
269 CTGCTGACCTCTTTATTTATCAAGTGTGAATGCGCGCGATGCTTCCACGCGGC 328
333 ATCGCAACCGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 392
329 TTAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 388
393 CTGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 452
389 GATGCGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 448
453 CAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 512
449 AACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 508
513 AACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 572
509 CCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
573 CCGAAGCGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 632
569 GTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
633 GTGCGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 692
629 TCCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 688
633 TC--GGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 750
669 CCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 745
751 CCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
746 TTAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 784
811 GTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 849

RESULT 2 805 bp DNA linear GSS 17-DEC-2002
B2556310
LOCUS pacel1-60_5617.x1 pacel1-60 Pseudomonas aeruginosa genomic clone
DEFINITION pacel1-60_5617, genomic survey sequence.

ACCESSION B2556310 GI:27166959
VERSION B2556310.1
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE Spencer,D.H., Raymond,C.K., Smith,B.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V. Whole-genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library J. Bacteriol. (2002) In press
JOURNAL Contact: Chris K. Raymond
COMMENT University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
Location/Qualifiers
1..805
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacel1-60_5617"
/clone.lib="pacel1-60"
/note="Clinical isolate 1-60 Whole genomic shotgun library."

ORIGIN

Query Match 6.7%; Score 194.8; DB 28; Length 805;
Best Local Similarity 60.6%; Pred.No. 2.5e-45;
Matches 319; Conservative 0; Mismatches 207; Indels 0; Gaps 0;
47 GACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 106
103 GATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 162
107 GTGCGTTCCTTTTGTATGATGATGATGATGATGATGATGATGATGATGATG 166
163 ATCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 222
167 TGAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 226
223 TGAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 282
227 GCGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 286
283 GCGGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
287 TTAGTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 346
343 ATCTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 402
347 GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 406
403 GCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 462
407 GCGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 466
463 GCGGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 522
467 CTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 526
523 CTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 582
527 CCGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 572
583 CCGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628


```

RESULT 3
BZ578932      853 bp      DNA      linear      GSS 17-DEC-2002
LOCUS      msh2_6070.y2 msh Pseudomonas aeruginosa genomic clone msh2_6070,
DEFINITION      genomic survey sequence.
ACCESSION      BZ578932
VERSION      BZ578932.1 GI:27213993
KEYWORDS      GSS.
SOURCE      Pseudomonas aeruginosa
ORGANISM      Pseudomonas aeruginosa
REFERENCE      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
AUTHORS      Pseudomonadaceae; Pseudomonas.
TITLES      1 (bases 1 to 853)
JOURNAL      Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
COMMENT      Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: Shotgun.
Location/Qualifiers
1..853
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="M5H"
/db_xref="taxon:287"
/clone="msh2_6070"
/clone_1lb="msh"
/note="Environmental isolate. Whole genomic shotgun
library."

ORIGIN
Query Match      6.6%; Score 191.8; DB 28; Length 853;
Best Local Similarity 60.0%; Pred. No. 2e-44;
Matches 336; Conservative 0; Mismatches 223; Indels 1; Gaps 1;

QY      29      CTGACGTTTCCCATATATGACCCCGTATTTGTTTCATCGCCGCTTACGGGTGCGCTGG 88
DB      168      CTGACGATCCCGACATCGATCCGTTGCGCTGCGCCATCGCCGCTGAAGATCACTGG 227
QY      89      TATGCGTTGATGATTTGTTGCGTTTCTTTTGTATGTTGTTGCAATCGCCGAGCG 148
DB      228      TAGGGGCTGATGATCGATCGGGATCGGGCGCCCTGCTGCGCTGCGCCCGGATG 287
QY      149      GATCGCCGCGGACGTGTTGACGCGGTGACCAAGTCTTCTGACTTGTATTTCGCGGCTTT 208
DB      288      AAGCGCTTCAACCCGACCTGACCAAGAGAGCGCTTCCACCTGCTTCTGCGTGGCG 347
QY      209      TTAGTGTAGTGAATGCGTGGCGGAGTTGTTATGATCTTCACAATTTGATCTGTC 268
DB      348      TGGGGCTGATCTCGGTGCGCCCTGCGCTACGTGTTCTTCAACCTGGAGCATGAC 407
QY      269      CTGCTGACCCCTTTATTTATTAAGATGTGACTGCGGCAATGCTCTTCCACGCGCG 328
DB      408      ATGCGCAACCCGACGCTGATCTTCAAGTCTGGAAGGGCGGCAATGCTTCCATGCGCG 467
QY      329      TTAATGGGTGATCAACCGGCAATGTTCTGTATGCGCGTAAACCAAGCAAGCACTTCTT 388
DB      468      CTGCTGCGCTGATGCTGCGCGTGTGATGCGCAAGCGCATGCGCAAGCACTTCTT 527
QY      389      GGTGTGCGCATTTTGTGCGCTTTAGTGCATTTGCGGTTTGGGATGGGACGATCGAT 448
DB      528      CAGCTGATGATCTCAATCGCCCGCTGCTGCGCATGCGCTGCGCGCGCGGCGCATCGCG 587
QY      449      AACTTATGATAGTGAATCTTGGGACAGATGACGATGCTGCTTGGGCTTTGTATTC 508

```

```

DB      588      AACTTATGATAGTGAATCTTGGGACAGATGATGCTGCTGCGCATGCTTCTT 647
QY      509      CCTATGTTGGCCCATCTGCGCGCATCTTTCACAGCTTTATGAATTCGCTTAGAAGCG 568
DB      648      CCCAAACG-CGCCCGCTGCGCGGCGCATCCCTGCAAGCTGTACAGATGCGCTGGAAGG 706
QY      569      GTGCTTCGTTCTTATCTCT 588
DB      707      CGTGGCTGTGCTATCTCT 726

RESULT 4
BZ559267      1141 bp      DNA      linear      GSS 17-DEC-2002
LOCUS      pacs2-164_1355.s1 pacs2-164 Pseudomonas aeruginosa genomic clone
DEFINITION      pacs2-164_1355, genomic survey sequence.
ACCESSION      BZ559267
VERSION      BZ559267.1 GI:27175318
KEYWORDS      GSS.
SOURCE      Pseudomonas aeruginosa
ORGANISM      Pseudomonas aeruginosa
REFERENCE      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
AUTHORS      Pseudomonadaceae; Pseudomonas.
TITLES      1 (bases 1 to 1141)
JOURNAL      Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
COMMENT      Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: Shotgun.
Location/Qualifiers
1..1141
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_1355"
/clone_1lb="pacs2-164"
/note="Clinical isolate 2-164 Whole genomic shotgun
library."

ORIGIN
Query Match      5.5%; Score 161; DB 28; Length 1141;
Best Local Similarity 64.6%; Pred. No. 2.7e-35;
Matches 272; Conservative 0; Mismatches 145; Indels 4; Gaps 2;

QY      380      ACCCTCTTGGTGGCGCATTTTGTGCGCCCTTATGAGCATTCGGTTGGGATGGA 439
DB      628      AGCTTCTTCAAGCTGATGAATTAATGCGCCCGCTGAGCCATGAGCTGGGCGCGG 569
QY      440      CGATCGTAACTTATGATGATGAATTTTGGGACAGTAAAGATGTCCTTGGGCT 499
DB      568      CGCATCGCACTTATCATCACTCGGAATGAGGGAAGGCAAGTCAAGATGCTCCCTGGGCC 509
QY      500      TTTGATTCCTATATGTTGGCCCATCGCGGCGCATCTTTCACAGCTTTATGATTTGCC 559
DB      508      ATGCTCTTCCCAACGCGGGGAC-GTGGCGGGATCTCTGACGCTGTACAGATTCGCC 450
QY      560      TTAGAGCGGTGTTCTGTTCTTATTTATTTATGTTTATGTTAAGCTGCTCCGCTA 619
DB      449      CTGGAAGCGTGGGGCTGTTCTATCTCTGCTGTTTCAACCGCAAGCCAGCGCCAGC 390
QY      620      GGCAGCGTATCGGACGTGTTTATGCTGATACGTTACATTCGCTTCTTGTGGAATAC 679
DB      389      GCTCGGTCTCGGCGCTGTTGCTGTGCTATGCGGATCTTCCGCTTGTGGAATTC 330

```

Qy 680 GTCCGTAGGACAGATGCTAGTTGGG---TCGTGTTGTTGGCTTCATTTCATAGGGGCA 736
 Db 329 GTCCGTAGGACAGATGCTAGTTGGG---TCGTGTTGTTGGCTTCATTTCATAGGGGCA 270
 Qy 737 ATCCCTCTTACCTATGATGATCATGCGATTTTATGATGTTGGTTCATCAAGGCC 796
 Db 269 GTCTCTGCTACCGATGATGCTGCGCGCATTTGCCCTGATGCTGCGCGCATCAAGGCC 210
 Qy 797 G 797
 Db 209 G 209

RESULT 5
 B2553447 1375 bp DNA linear GSS 17-DEC-2002
 LOCUS pasci-60_4149, genomic survey sequence.
 DEFINITION
 ACCESSION B2553447
 VERSION B2553447.1 GI:27159498
 KEYWORDS GSS.
 SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 1375)
 AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.R., Hastings,M.,
 Burns,J.L., Kaul,R. and Olsen,M.V.
 TITLE Whole-Genome-Sequence Variation among multiple isolates of
 Pseudomonas aeruginosa library
 JOURNAL J. Bacteriol. (2002) In press
 COMMENT Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.

FEATURES
 source Location/Qualifiers
 1..1375
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="1-60"
 /db_xref="taxon:287"
 /clone="pasci-60_4149"
 /clone_lib="pasci-60"
 /note="clinical isolate 1-60 whole genomic shotgun
 library."

ORIGIN

Query Match 5.2%; Score 152; DB 28; Length 1375;
 Best Local Similarity 57.7%; Pred. No. 1.3e-32;
 Matches 291; Conservative 0; Mismatches 210; Indels 3; Gaps 1;
 Qy 270 TTGCTGACCTCTTATTATTATCAAGTGTGAGTGGCGGATGCTCTTCACGGCGCT 329
 Db 84 TTCGCGACGCCCGGTGATTTTCATGCTGTGAAAGGCGCATGCTCCATGCGCGGC 143
 Qy 330 TATTGGGTGTATACACGCCCATGTTCTGTATGCGCGCAAAACCAACGACCTCTTTG 389
 Db 144 TGCTGGGGGTATGCTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 203
 Qy 390 GTGTGGCGATTTTGTGCTCCCTTATGTCATTCGTTTGGGATGGGATCGTATCGGTA 449
 Db 204 AGCTATGATCTTATCGCCCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 263
 Qy 450 ACTTATGATATGTAACCTTTTGGGACAGATACGATGTGCTTGGCTTTTATTTCC 509
 Db 264 ACTTATGATATGTAACCTTTTGGGACAGATACGATGTGCTTGGCTTTTATTTCC 323
 Qy 510 CTAATGTGGCCCATGCGCGCGCATCTTCACAGCTTATGAATTCGCTTAGAAGGCG 569

Db 324 CCAAGCGGCGCCGCTGCGCGGACATCCCGAAGTTGATACAGTTGCGCCCTGAAGGCC 383
 Qy 570 TGGTTCGTTCTTATTCTTAATGTTTATTTGTTAATGTTAAGCTGTCGCGCTAGGAGGTA 623
 Db 384 TGGCATGTTCTGATCTCTGCGTGGCCACCCGATGCGACGATGACCGCTTGTCT 443
 Qy 630 CCGGATCTTTTATGCTGATACGATACGATACGATACGATACGATACGATACGATACG 689
 Db 444 CTGGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 503
 Qy 690 CAGATGCTAGTTGG---GTCTGTTTGGTGGCTTCATTTCATAGGGGCAATCTCTCT 745
 Db 504 CGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 563
 Qy 747 TACCTATGATATCATCGTATTT 770
 Db 564 TACCGTGAATGCTGCTTCT 587

RESULT 6
 U83227 406 bp DNA linear GSS 21-FEB-2001
 LOCUS U83227
 DEFINITION
 ACCESSION U83227
 VERSION U83227.1 GI:1786113
 KEYWORDS GSS.
 SOURCE Pseudomonas putida
 ORGANISM Pseudomonas putida
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 406)
 AUTHORS Caspi,R., Tebo,B.M. and Haygood,M.G.
 TITLE c-type cytochromes and manganese oxidation in Pseudomonas putida
 JOURNAL Appl. Environ. Microbiol. 64 (10), 3549-3555 (1998)
 MEDLINE 98432765
 PUBMED 9758766
 COMMENT Contact: Caspi R

FEATURES
 source Location/Qualifiers
 1..406
 /organism="Pseudomonas putida"
 /mol_type="genomic DNA"
 /strain="MnB1; ATCC 23483"
 /db_xref="taxon:303"
 /clone="transposon-tagged mutant UT502"
 /clone_lib="Pseudomonas putida MnB1; ATCC 23483"
 /note="manganese oxidizing bacterium"
 Class: transposon-tagged.

ORIGIN

Query Match 4.0%; Score 115.6; DB 29; Length 406;
 Best Local Similarity 59.8%; Pred. No. 3.8e-22;
 Matches 224; Conservative 0; Mismatches 145; Indels 12; Gaps 2;
 Qy 399 ATTTGTTGCTCCCTTATGTCATTCGTTTGGGATGGGATCGTATGCTTATGTA 458
 Db 15 ATTTGTTGCTCCCTTATGTCATTCGTTTGGGATGGGATCGTATGCTTATGTA 74
 Qy 459 ATAGTGAATTTTGGGACAGATGATGCTTGGGCTTTTGTATTTCC----- 510
 Db 75 AGCCGAGCTGTGGGCAAAACCAACGATGTCATGCGGATGATCTTCGCGCTTCA 134
 Qy 511 -TAATGTGGCCCATGCGCGCGCATCTTCACAGCTTATGAATTCGCTTAGAAGGCG 569

Db 135 GCGACCCAGCCAGCTCCACGACCCGTCGAGCTGATCCAGTTGCGCCTGGAGGGG 194
 QY 570 TGGTGTCTTATCTTAATGGTTATGTGTAACCTGTCGCTGAGGAGGAT 629
 Db 195 TGGACATGTTGATCTCTGCTGTTTCGCGCAAGCCGCGCAGCATT 254
 QY 630 CCGGAGCTTTTATAGTGAATAGGTAACCTGCTTCCTTGAATAGCTCGTAGC 689
 Db 255 CGGACATGTTGCGCTGTTCTAGCGCATCTTCGCTTCATGTCGAAATGTCGGGGTGC 314
 QY 690 CAGATGCTGAGTTGGG---TCTGTTGGTGGCTTCATTTCAATGGGGCAATCTCTCT 746
 Db 315 CGATGCGCAGCTGGGCTATATGCGCTGGGGCTGGCTGACCATGAGTCAATTCGTGTG 374
 QY 747 TACCTATGATGATCATGCTATTTTATGATGAT 777
 Db 375 TGCCGATGATCTGCTGGCTGGCCTGAT 405

RESULT 7
 B2561697 1034 bp DNA linear GSS 17-DEC-2002
 LOCUS pacs2-164_3389.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
 DEFINITION pacs2-164_3389, genomic survey sequence.
 ACCESSION B2561697
 VERSION B2561697.1 GI:27182207
 KEYWORDS GSS.
 SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
 1 (bases 1 to 1034)
 Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
 Burns, J.L., Kaul, R. and Olsen, M.V.
 Whole-Genome-Sequence Variation among Multiple Isolates of
 Pseudomonas aeruginosa Library
 J. Bacteriol. (2002) In press
 Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: shotgun.
 Location/Qualifiers
 source 1..1034
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="2-164"
 /db_xref="taxon:287"
 /clone="pacs2-164_3389"
 /clone_1id="pacs2-164"
 /note="clinical isolate 2-164 whole genomic shotgun library."

Query Match 3.8%; Score 111.2; DB 28; Length 1034;
 Best Local Similarity 53.9%; Pred. No. 1.2e-20;
 Matches 342; Conservative 0; Mismatches 278; Indels 14; Gaps 5;

QY 29 CTGACGTTTCCCATATGACCCCGTATGTTTCGATCGGCCCTTAGCGGTGCGCTG 88
 Db 370 CTGACGATCCCGACGATGATCCGCTGCGCTGCCATCGCCCGCTGAAGATCACTGG 429
 QY 89 TATGAGCTGATGATTTGGTGGGTTTCTTTTTCATGATGATGTTGGCAATGCCGAGG 148
 Db 430 TACGGGCTGATGATCTGATCGATCGGATCGGGGCGCTGCTGTGCGTGGCCCGGATG 489
 QY 149 GATCGCGCGGAGATGTTGAGCGCGTGAAGATCTCTGACTTGTATTC--GCCGCG 205
 Db 490 AAGGCTTGCACCGACCTGACCAAGAGAGCGCTTCCGACTGTGCTTGGGTTCCG 549

QY 206 TTTTAGGTAGTATGATCGTGGCCGAGTTGGTATGATCTTCTACATTTTATGATCG 265
 Db 550 CTGGGGGGGATACCTCCGATGGCGGCTGGGAGTACGATGCTGTCTACAACTGGAGAG 609
 QY 266 TTCCTGCTGACCCCTTTTATTTATTCAAAGTGTGAGATGGCGGATCTCTCCAGCGC 325
 Db 610 TACATGCGCAACCCACGCTGATCTTCAGGTCGAGGAGGCGGATCTCTTCATGATG 669
 QY 326 GCGTATTTGGGTGATCACCGCCATGTTCTGTATGTCGCGCTTAAACCAAGCAGCTTC 385
 Db 670 GGGCTGCTCGGATATGCTGCGGTCTGATGTTGGTTGGCAGAGCGCATGGCAAGAGC 729
 QY 386 TTGGTGTGGCCGA----TTTGTGGCCCTTATAGTCCATTCGTTTGGGGATGGAC 440
 Db 730 TTCCTTACGCTGAGGAGACTTCATTCGCCCCGTTGGTCCCATGGGCTGGGCGGCGGC 789
 QY 441 GTATGCGTACTTTATGAAATGTAATTTGGGAGCAG---TAACGATGTCCTTGGG 497
 Db 790 GCATTGGGAATTTTATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 849
 QY 498 CTTTGTATTCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 556
 Db 850 CAGGAGCTTTCCCAAGCTCGGCCCCCTTCGCGGCGATCCCTTACAAATGTCGCAATTC 509
 QY 557 GCCTTAGAAGGCGGTGCTCTGTTCTTTA--TTCTTAATGTTTATTTGTAACCTGTC 614
 Db 910 GCCTTAGAAGGCGGCGGCGCGGCTTTTAAATCTTGGGTTGTTAACCACCAAGCCGGGC 969
 QY 615 CGTAGGACGATGATCGGAGCTGTTTATGCTGG 648
 Db 970 CGCCGCGCTTGCTTCCGCTCTTTCCGCGCTTG 1003

RESULT 8
 CD379073 878 bp mRNA linear EST 31-MAY-2003
 LOCUS PTM004242 Paeodactylum tricornutum Uni-Zap XR Paeodactylum
 tricornutum cDNA 5', mRNA sequence.
 DEFINITION
 tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD379073
 VERSION CD379073.1 GI:31254687
 KEYWORDS EST.
 SOURCE Paeodactylum tricornutum
 ORGANISM Paeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Paeodactylaceae; Paeodactylum.
 1 (bases 1 to 878)
 Scala, S., Carels, N., Falciatore, A., Chiusano, M.L. and Bowler, C.
 Genome properties of the diatom Paeodactylum tricornutum
 Plant Physiol. 129 (3), 993-1002 (2002)
 12111123
 12114555
 CONTACT: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.szn.it
 Diatom EST Database (http://avesthagen.sznbowler.com)
 Seq primer: T3 backward
 POLYA=yes.

FEATURES
 source 1..878
 Location/Qualifiers
 /organism="Paeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_1id="Paeodactylum tricornutum Uni-Zap XR"
 /note="vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
 Xho I"

Query Match 3.8%; Score 111; DB 14; Length 878;
 Best Local Similarity 59.8%; Pred. No. 1.2e-20;
 Matches 186; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 2474 ATGATTTCTAGAGGTAGAGATCAGACCGCGAGAAAGATTGAAAAATTCGTGATGTTG 2533
 DB 332 ATGACTTCTACAGAGAAAGATGATGGCCACCGACAGAAATTAACATATGTTGAGAGGTTG 273
 QY 2534 GCACGTGGCGAAGACCGCGCAAACTCTTTTACAGGGCATATGGCCAGAGGCTGACGAGA 2593
 DB 272 GCACGACGAGAAAGACCGCTGAGCTGATCCCATCGCCCATTTGGCCCATAGGAGATTT 213
 QY 2594 GTATTGGTCAAGTAGTACCAATATTGGACCCATCAACATAGAGAAATCGGGTTTCAAC 2653
 DB 212 GTGTGGTGAAGTGAAGTACCAATATTCGCTCCCATGATCATGTAATTCCTTGGCGGAG 153
 QY 2554 GGTAAACCAACCGGCAAGACACCAAAATATAGAACTGACCGTGTGAGATTGAATG 2713
 DB 152 GAGACAGCGGATTCGACCAAGACACCAATGAGTCTTTGAGAGAAAGCTGAGC 93
 QY 2714 AGTCCGTGGCAGTAAACCAATCATCATCTGCAATTTGGTGGAGCAAAATTCAAAT 2773
 DB 92 AATACAGTCGAGAAATACCAACCATTAAGACGACATGGATTCGTGTGCGCGAGC 33
 QY 2774 AGAATTGTGC 2784
 DB 32 AACTTCGCGC 22

RESULT 9
 CD382942/c 921 bp mRNA linear EST 31-MAY-2003
 LOCUS PTM08112 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 DEFINITION tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD382942
 VERSION CD382942.1 GI:31258556
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 1 (bases 1 to 921)
 REFERENCE Scala, S., Carels, N., Falciatore, A., Chiusano, M.L. and Bowler, C.
 Genomic properties of the diatom Phaeodactylum tricornutum
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 2211123
 PUBMED 12114555
 COMMENT Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.szn.it
 Diatom EST Database (http://avesthagen.sznbowler.com)
 Seq primer: T3 backward
 POLYA=yes.

FEATURES
 source location/Qualifiers
 1..921
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
 Xho I"

ORIGIN
 Query Match 3.8%; Score 111; DB 14; Length 921;
 Best Local Similarity 59.8%; Pred. No. 1.3e-20;
 Matches 186; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 2474 ATGATTTCTAGAGGTAGAGATCAGACCGCGAGAAAGATTGAAAAATTCGTGATGTTG 2533

DB 353 ATGACTTCTACAGAGAAAGATGATGGCCACCGACAGAAATTAACATATGTTGAGAGGTTG 294
 QY 2534 GCACGTGGCGAAGACCGCGCAAACTCTTTTACAGGGCATATGGCCAGAGGCTGACGAGA 2593
 DB 293 GCACGACGAGAAAGACCGCTGAGCTGATCCCATCGCCCATTTGGCCCATAGGAGCATT 234
 QY 2594 GTATTGGTCAAGTAGTACCAATATTGGACCCATCAACATAGAGAAATCGGGTTTCAAC 2653
 DB 233 GTGTGGTGAAGTGAAGTACCAATATTCGCTCCCATGATCATGTAATTCCTTGGCGGAG 174
 QY 2654 GGTAAACCAACCGGCAAGACACCAAAATATAGAACTGACCGCTTGAAGATTGAATC 2713
 DB 173 GAGACAGCGGATTCGACCAAGACACCAATTCAGAGTCTTTGAGAGAAAGACTGAGC 114
 QY 2714 AGTCCGTGGCAGTAAACCAATCATCATCTGCAATTTGGTGGAGCAAAATTCAAAT 2773
 DB 113 AATACAGTCGAGAAATACCAACCATTAAGACGACATGGATTCGTGTGCGCGAGC 54
 QY 2774 AGAATTGTGC 2784
 DB 53 AACTTCGCGC 43

RESULT 10
 CD376447/c 855 bp mRNA linear EST 31-MAY-2003
 LOCUS PTM01614 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 DEFINITION tricornutum cDNA 5', mRNA sequence.
 ACCESSION CD376447
 VERSION CD376447.1 GI:31252061
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum
 Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.
 1 (bases 1 to 855)
 REFERENCE Scala, S., Carels, N., Falciatore, A., Chiusano, M.L. and Bowler, C.
 Genomic properties of the diatom Phaeodactylum tricornutum
 JOURNAL Plant Physiol. 129 (3), 993-1002 (2002)
 MEDLINE 2211123
 PUBMED 12114555
 COMMENT Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.szn.it
 Diatom EST Database (http://avesthagen.sznbowler.com)
 Seq primer: T3 backward
 POLYA=yes.

FEATURES
 source location/Qualifiers
 1..855
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
 Xho I"

ORIGIN
 Query Match 3.7%; Score 106.2; DB 14; Length 855;
 Best Local Similarity 58.8%; Pred. No. 3.1e-19;
 Matches 183; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 2474 ATGATTTCTAGAGGTAGAGATCAGACCGCGAGAAAGATTGAAAAATTCGTGATGTTG 2533
 DB 323 ATGACTTCTACAGAGAAAGATGATGGCCACCGACAGAAATTAACATATTCGTGAGAGGTTG 264
 QY 2534 GCACGTGGCGAAGACCGCGCAAACTCTTTTACAGGGCATATGGCCAGAGGCTGACGAGA 2593

Db 263 GCACGACGAAAGACGCTGAGTTGATCCCATTCGCCATTGGCCCATAGGAGATG 204

QY 2594 GTATTGGTCAAGTAGTACCAATATTGGACCCATCCATAGGAATCGCGTTTCAACC 2653

Db 203 GTGTGGTGAAGTGAAGTACCAATATTGGCTCCCATGATCATGTAAATACCTTGGCGGAGC 144

QY 2654 GGTAAACCCACGGGCAAGACCAACCAATATTAAGTACCGGTGTGAGATTGATC 2713

Db 143 GAGCAGCGGATCCGACCAAGACACACATGAGTATTGTTGAAGCAGCTGAAGC 84

QY 2714 AGTCCCTTCGCACTTAACCAATCATCATCTGCAATTTGGTGGAGCAATTTCAAT 2773

Db 83 AATACATTCGCAAGAAATACCAACCATTAAGACGAGGTGGATTCGTGTGCGCGAAGC 24

QY 2774 AGAATTTGGC 2784

Db 23 AACTCTCCGGC 13

RESULT 11 843 bp mRNA linear EST 19-NOV-2002

CA588286/hab49a11.y1 Fugu UT6 adult gut Takifugu rubripes cDNA clone IMAGE:6351812 5' similar to TR:Q91237 Q91237 NM/PI COTRANSPORT SYSTEM PROTEIN. ; mRNA sequence.

ACCESSION CA588286

VERSION CA588286.1 GI:25132864

KEYWORDS EST.

SOURCE Takifugu rubripes (Fugu rubripes)

ORGANISM Takifugu rubripes

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Takifugu.

1 (bases 1 to 843)

Clark, M., Johnson, S.L., Lebrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.

Washu Zebrafish EST Project 1998

Unpublished (1998)

CONTACT: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrafish@wustl.edu

Library materials provided by G. Elgar (UK MRC HGMP-RC) Library constructed by Drs. K. Kawakami, M. Sasaki, S. Sugano, K. Kikuchi and S. Watabe (University of Tokyo, Institute of Medical Science and Laboratory of Aquatic Molecular Biology and Biotechnology) DNA Sequencing by: Washington University Genome Sequencing Center Clone through the I.M.A.G.E. Consortium/LNL, send email to: info@image.lnl.gov

Seq primer: T3 ET from Amersham

High quality sequence stop: 496.

FEATURES

Source

1..843

Location/Qualifiers

organism="Takifugu rubripes"

molecule="mRNA"

db_xref="taxon:31033"

clone="IMAGE:6351812"

tissue_type="gut"

dev_stage="adult"

lab_host="DH10B (T1 phage-resistant)"

clone_lib="Fugu UT6 adult gut"

note="Vector: pBluescript-FL; Site 1: PflM I (CCANNNTGG); Site 2: PflM I (CCANNNTGG); BamHI-SmaI sites were converted to BamHI-PflM-SfiI-PflM sites (SmaI is destroyed). Other part of the vector is untouched. The

ORIGIN

Query Match 3.6%; Score 104; DB 14; Length 843;

Best Local Similarity 58.9%; Fred. No. 1.4e-18;

Matches 198; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

2427 GCAGAGCGAAACGACGCGAGTGTAGACTTTCTTAATATGCCAAACATCATTTTGAAG 2486

Db 721 GGAAGAGCTATATCAGAACTGTGGAAGCTTTGAAGAACTCTGTACACCTCCAG 662

QY 2487 GTAGAAGATCAGACCGCGAGAGAGATTGAAAAATCTGTGATGTGCACTGGCGAAG 2546

Db 661 GCAGAAGATCAACACGACAGCCAGTTAAAGTCTGTGACGCTCCGCCAGCAAGG 602

QY 2547 CAGGCGGAAACTCTCTTTACAGCGCATATGAGCGAGCTGACGAGATTGTGCACAG 2606

Db 601 CCTGGGAACTCTGTTGATCTCCCGCTGATCATGTGCCAGATGTGTTGGTGAAG 542

QY 2607 TAGTACCAATATTGGACCCCATACCATAGGAATGCGGTTT--CAACCGTAAACCAAC 2663

Db 541 AGTCCCGATGTTGGCGCCCATGATATCGGACCGCGGTGACATCATCATATTCAG 482

QY 2664 CGGCAACGACCAACCAATATAGAGACCGCTCTGAGATGTGATGTCGCGCTTG 2723

Db 481 AGGACACCATGTGTGCCCAAGAGGAGAGATCTGAAGCTTGACGACGACGTTGA 422

QY 2724 CCAGTAAACCATCATCATCTGCAATTTGGTGG 2759

Db 421 CTAAACCCCGATCAACGCCACGACGAGGTTGG 386

RESULT 12 813 bp mRNA linear EST 19-NOV-2002

CA588344/hab50c01.y1 Fugu UT6 adult gut Takifugu rubripes cDNA clone IMAGE:6351745 5' similar to TR:Q91237 Q91237 NM/PI COTRANSPORT SYSTEM PROTEIN. ; mRNA sequence.

ACCESSION CA588344

VERSION CA588344.1 GI:25132922

KEYWORDS EST.

SOURCE Takifugu rubripes (Fugu rubripes)

ORGANISM Takifugu rubripes

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Takifugu.

1 (bases 1 to 813)

Clark, M., Johnson, S.L., Lebrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.

Washu Zebrafish EST Project 1998

Unpublished (1998)

CONTACT: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrafish@wustl.edu

Library materials provided by G. Elgar (UK MRC HGMP-RC) Library constructed by Drs. K. Kawakami, M. Sasaki, S. Sugano, K. Kikuchi and S. Watabe (University of Tokyo, Institute of Medical Science and Laboratory of Aquatic Molecular Biology and Biotechnology) DNA Sequencing by: Washington University Genome Sequencing Center Clone

distribution: Fugu clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN, send email to: info@image.llnl.gov
 Seg primer: T3 RT from Amersham
 High quality sequence stop: 497.
 Location/Qualifiers

FEATURES

1..813

/organism="Takifugu rubripes"
 /mol_type="mRNA"
 /db_xref="taxon:31033"
 /clone="IMAGE:6351745"
 /rname_type="gut"
 /dev_stage="adult"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_id="Fugu UT6 adult gut"
 /note="Vector: pBluescript-FL; Site 1: PflM I (CCANNNTGG); Site 2: PflM I (CCANNNTGG); BamH-SmaI sites were converted to BamH-PflM-Sfi-PflM sites (SmaI is destroyed). Other part of the vector is untouched. The cDNA is inserted between two PflM sites in T3 (5') to T7 (3') direction. Library materials provided by G. Elgar (UK MRC H&MP-RC) and constructed and donated by Drs. K. Kawakami, M. Sasaki, S. Sugano, K. Kikuchi and S. Warabe (University of Tokyo, Institute of Medical Science and Laboratory of Aquatic Molecular Biology and Biotechnology)."

ORIGIN

Query Match 3.6%; Score 103.8; DB 14; Length 813;
 Best Local Similarity 59.2%; Pred. No. 1.6e-18;
 Matches 196; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

2432 GGGCAACGAGCCAGTGTGAGACTTCTCTAAATGCGCAACATCATTTCTAGAGAGG 2491
 713 GACGTATCGAAGACTGGGTGAAGTCTGAAAGAACTCTGTAGCCACCTCCAGAGGAGA 654
 2492 AAGATCGACACCGGAGAAATTGAAATTCGTGATGTGTGCACTGGCGAAAGACCG 2551
 653 AGATCAACACCGACGACAGTAAAGAGTCGTGAGCGGTCGCCAGAAAGCCCTG 594
 2552 CGAACTCTTTTACAGCGCATATGGCCAGGCTGACGAGATTTGTTCACAGTACTA 2611
 593 CGGAACCTGTTTCATCCCCCGCTGCATCATGCGCAGATGGTGTGGACAGAGTTC 534
 2612 CCAATATTGACACCATCATCAGATGCGCGTTT---CAACCGGTAACCCACCGGCA 2668
 533 CGGATGTGGCGCCCATATATATGCGACCCCGCTTCGACATCAGTATTTCCAGAGGAC 474
 2669 ACGAGACCAACATATATGAAAGTCAACCGTCTTGAGATTAATCAAGTCCGTCCT 2728
 473 ACCATGTGACCAATAATGAGAGAGAGTGTGAGCTTGACAGACACGATGACTAAC 414
 2729 AAACCAATCATCATCCGCGCAATTGGGTGGG 2759
 413 ACCCGATCACCAAGCCCAAGCGGTTGG 383

RESULT 13
 CA475083 828 bp mRNA linear EST 12-NOV-2002
 LOCUS CA475083
 DEFINITION AGENCOURT 10670196 NCI CGAP_ZKId1 Danio rerio cDNA clone
 IMAGE:6797422 5', mRNA sequence.
 ACCESSION CA475083
 VERSION CA475083.1 GI:24931435
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
 REFERENCE 1 (bases 1 to 828)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Leonard I. Zon, M.D.
 cDNA Library Preparation: Invitrogen Corp
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
 Plate: LLM14310 row: K column: 21
 High quality sequence stop: 433.
 Location/Qualifiers

FEATURES

1..828

/organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:6797422"
 /lab_host="DH10B (T1-resistant)"
 /clone_id="NCI CGAP_ZKId1"
 /note="Organ: kidney; Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: 0130 dt. Average insert size 1.8 kb. Constructed by J. Wang (Research Genetics, Invitrogen Corp) from tissue donated by L. Zon (Harvard University). Note: this is a NCI CGAP library."

ORIGIN

Query Match 3.5%; Score 103.2; DB 14; Length 828;
 Best Local Similarity 60.5%; Pred. No. 2.4e-18;
 Matches 187; Conservative 0; Mismatches 119; Indels 3; Gaps 1;

2465 ATGCCAACATCATTTCTAGAGGTAGGAAGTACGACCGCGAGAGATTGAAATATCG 2524
 521 ATGCTGTGATGCTTTCCAGCGGACGACGACGACGACTACACCACTTAACATATCG 462
 2525 TGGATGTGGCACTGGCGAAGACACGCGAACTCTTTTACAGCCCATATGSCCAAGG 2584
 461 TGCATGTGGGACACGACGAAAGCTGTTGAACTCTCTCTCTCTCTCTCTCTCTCTCT 402
 2585 CTGACGAGATTTGGTGTACAGTGTACCATATTTGGACCATCATCAGATGAGATCGCG 2644
 401 GCCACGATGTGTGTGTACAGATGTGCAATGTTTGAAGCCCATGATGATGGGATGGCG 342
 2645 GTT---TCAACCGGTAAACCCACCGGCAAGAGACCAATATATGAGTACCGTCTT 2701
 341 GATCCCATCATCAGCAATTCAGAGAAACGACGATGACCAATGAGATGAGTCTCTG 282
 2702 GAGATTTGATATGTCGCGTGTCCACTTAACCATCATCATCTCTGCAATTGGGTGGAA 2761
 281 GAACTTTGACAGAGACGATACCAATTTCTCCACCACTCTCCGACAGGGGTAGAA 222
 2762 GCAATTTCA 2770
 221 AGAACCGCA 213

RESULT 14
 BM039803 685 bp mRNA linear EST 19-OCT-2002
 LOCUS BM039803
 DEFINITION BM039803 Nori Satoh unpublished cDNA library, blood cells Ciona intestinalis cDNA clone c1bd041k19 5', mRNA sequence.
 ACCESSION BM039803
 VERSION BM039803.1 GI:24140498
 KEYWORDS EST.
 SOURCE Ciona intestinalis
 ORGANISM Ciona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cloniidae; Ciona.
 REFERENCE 1 (bases 1 to 685)
 AUTHORS Satoh, Y., Satake, M., Azumi, K., Nomura, M., Shin-i, T., Kohara, Y. and Satoh, N.
 TITLE Expressed genes in Ciona intestinalis (2002)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 15:39:05 / Search time 888 Seconds
(without alignments)

16062.201 Million cell updates/sec

Title: US-09-700-712a-1

Perfect score: 2909
Sequence: 1 gagaagttgttgcctc.....cagcgaatcgcgtcag 2909

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	112.6	38.2	1149	US-10-282-122A-40937	Sequence 40937, A
2	842.4	29.0	852	US-10-282-122A-40936	Sequence 40936, A
3	744.2	25.6	1830121	US-10-329-960-1	Sequence 1, Appl
4	744.2	25.6	1830121	US-10-329-960-1	Sequence 1, Appl
5	505.4	17.4	852	US-10-282-122A-30692	Sequence 30692, A
6	462.2	15.9	852	US-10-282-122A-22162	Sequence 22162, A
7	403.8	13.9	849	US-10-282-122A-32325	Sequence 32325, A
8	248.2	8.5	876	US-09-741-669-182	Sequence 182, App
9	152.2	5.2	640681	US-09-790-988-1	Sequence 1, Appl
10	142.2	4.9	96109	US-10-672-787-35	Sequence 35, Appl
11	115.4	4.0	2731748	US-10-297-465A-1	Sequence 1, Appl
12	84	2.9	2440	US-10-152-119A-1589	Sequence 1589, Ap
13	77.6	2.7	3950	US-10-152-119A-1588	Sequence 1588, Ap
14	75.4	2.6	2208	US-10-276-774-1331	Sequence 1331, Ap

15	75.4	2.6	2280	16	US-10-173-999-149	Sequence 149, App
16	75.4	2.6	3952	13	US-10-116-802-36	Sequence 36, Appl
17	75.4	2.6	4137	14	US-10-052-664-2	Sequence 2, Appl1
18	73.8	2.5	4111	13	US-10-257-021-146	Sequence 146, Appl
19	73.8	2.5	4135	15	US-10-097-140-220	Sequence 220, App
20	72.2	2.5	687	17	US-10-404-460-258	Sequence 258, App
21	70.2	2.4	1800	13	US-10-169-395-71	Sequence 71, Appl
22	70.2	2.4	2016	13	US-10-169-395-71	Sequence 81, Appl
23	70.2	2.4	2124	16	US-10-108-260A-1011	Sequence 1011, Ap
24	70	2.4	1164	9	US-09-738-626-5017	Sequence 3017, Ap
25	70	2.4	1333	13	US-10-627-476-511	Sequence 511, App
26	70	2.4	3309400	9	US-09-738-626-511	Sequence 1, Appl1
27	69.2	2.4	969	13	US-10-282-122A-33224	Sequence 33224, A
28	68.2	2.3	795	13	US-10-282-122A-42178	Sequence 42178, A
29	67.6	2.3	969	13	US-10-282-122A-14737	Sequence 14737, A
30	66.4	2.3	792	13	US-10-282-122A-12730	Sequence 12730, A
31	65.4	2.2	722	17	US-10-332-859-254	Sequence 254, App
32	64.8	2.2	1478	15	US-10-084-817-332	Sequence 332, App
33	64.8	2.2	1536	11	US-09-877-178-11	Sequence 11, Appl
34	64.8	2.2	1536	13	US-10-342-887-556	Sequence 556, App
35	64.8	2.2	1536	13	US-10-172-118-556	Sequence 556, App
36	64.8	2.2	1536	13	US-10-388-560-350	Sequence 350, App
37	64.8	2.2	1536	16	US-10-159-563-244	Sequence 244, App
38	64.8	2.2	1536	17	US-10-283-975A-254	Sequence 254, App
39	64.8	2.2	3298	15	US-10-240-965-200	Sequence 200, App
40	64.6	2.2	795	9	US-09-741-669-181	Sequence 181, App
41	64.6	2.2	795	13	US-10-282-122A-6574	Sequence 6574, Ap
42	64.6	2.2	798	9	US-09-738-626-942	Sequence 942, App
43	64.6	2.2	1200	43	US-09-954-197-1	Sequence 1, Appl1
44	64.6	2.2	2730	9	US-09-987-446-1	Sequence 1, Appl1
45	64.4	2.2	969	13	US-10-282-122A-14234	Sequence 14234, A

ALIGNMENTS

RESULT 1
US-10-282-122A-40937/c
Sequence 40937, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 40937
 LENGTH: 1149
 TYPE: DNA
 ORGANISM: *Vibrio cholerae*
 US-10-282-122A-40937

Query Match 38.2%; Score 1112.6; DB 13; Length 1149;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1118; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

1783 TTAAGCGGTGAGGCAAGATGCTGCGGGATGACGCAAAACACCCCAATAGTACT 1842
 1149 TTAAGCGGTGAGGCAAGATGCTGCGGGATGACGCAAAACACCCCAATAGTACT 1090
 1843 CACCACACCACTTTTGTCTTACAGCCCAAGTTAGATGAGCTCAGCACTTTAATAG 1902
 1089 CACACACCACTTTTGTCTTACAGCCCAAGTTAGATGAGCTCAGCACTTTAATAG 1030
 1903 CAGTTCCGCTAAGAAAGAAATCCGTAAATCAGACCGCTGCGCATCAAGTTAAAGCTTAA 1962
 1029 CAGTTCCGCTAAGAAAGAAATCCGTAAATCAGACCGCTGCGCATCAAGTTAAAGCTTAA 970
 1963 GTGACACGACCAATTTTGCAGAGCAACACGCAAACTCAGCAAGACCGGTTGCGG 2022
 969 GTGACACGACCAATTTTGCAGAGCAACACGCAAACTCAGCAAGACCGGTTGCGG 910
 2023 GAGCAGACAGTAATACAGTCCCATGTTGCGCACTTAAGTAATGGTGAATTTACG 2082
 909 GAGCAGACAGTAATACAGTCCCATGTTGCGCACTTAAGTAATGGTGAATTTACG 850
 2083 CACTTTGAGACGCGCAGACGCCACGAGAGAACCATTAAGCTGTTGCTGATGATGA 2142
 849 CACTTTGAGACGCGCAGACGCCACGAGAGAACCATTAAGCTGTTGCTGATGATGA 790
 2143 TTGAATTAATCCGTAAACACTGTACCTGAAGCAATACCGTGAAGTGGGCTTGCGCAAT 2202
 789 TTGAATTAATCCGTAAACACTGTACCTGAAGCAATACCGTGAAGTGGGCTTGCGCAAT 730
 2203 CGCATTTTGTGAATTTTCACTGCGGCGGCAACCAATCACTTTCACTGATGATGATGATG 2262
 729 CGCATTTTGTGAATTTTCACTGCGGCGGCAACCAATCACTTTCACTGATGATGATGATG 670
 2263 CACCGTAATGCGCAGAAATGCTGCGATACCCCAATAGATAGTGGCAACCAACGAA 2322
 669 CACCGTAATGCGCAGAAATGCTGCGATACCCCAATAGATAGTGGCAACCAACGAA 610
 2323 AGTATTACCAATACCGAAAGCTGGGTTTCAAGCCTGTGATGACAGGTTTGGTAATCG 2382
 609 AGTATTACCAATACCGAAAGCTGGGTTTCAAGCCTGTGATGACAGGTTTGGTAATCG 550
 2383 TTATATTAATCAAACTTTTCACTGATGATGATGATGATGATGATGATGATGATGATG 2442
 549 TTATATTAATCAAACTTTTCACTGATGATGATGATGATGATGATGATGATGATGATG 490
 2443 CCAGTGTGAGACTTTTCTTAAATGCAAAATCATATTTCTAAGAGTGAAGATCGAC 2502
 489 CCAGTGTGAGACTTTTCTTAAATGCAAAATCATATTTCTAAGAGTGAAGATCGAC 430
 2503 CCGGAGAGAGTTGAAAAATCGTGAATGCGGCACTGGCGGAAAGCAACCGGAAATCTTTC 2562
 429 CCGGAGAGAGTTGAAAAATCGTGAATGCGGCACTGGCGGAAAGCAACCGGAAATCTTTC 370
 2563 TTATACAGCGATATGCGCAAGGCTGACGAGATATGCTACAGTATGATGATGATGATG 2622
 369 TTATACAGCGATATGCGCAAGGCTGACGAGATATGCTACAGTATGATGATGATGATG 310

2623 ACCCATCAACATAGAAATGCGGTTTCAACCGGTAAACCCAGGCAACGAGACCAACAT 2682
 309 ACCCATCAACATAGAAATGCGGTTTCAACCGGTAAACCCAGGCAACGAGACCAACAT 250
 2683 AATAGAGTCAACCGGTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 2742
 249 AATAGAGTCAACCGGTTTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 190
 2743 TCTGTCAATGGGTGGAAGCAAAATTCAAATAGAACTTTGCTGATGCGGTTGCGCA 2802
 189 TCTGTCAATGGGTGGAAGCAAAATTCAAATAGAACTTTGCTGATGCGGTTGCGCA 130
 2803 TTTAAACCGGCTGCGCAACATGCGACTGCAAGAGTATGATTAATACGATGAAGCCAA 2862
 129 TTTAAACCGGCTGCGCAACATGCGACTGCAAGAGTATGATTAATACGATGAAGCCAA 70
 2863 GTTTCGCAACGTAAGACCTTTCGATGATGATGATGATGATGATGATGATGATGATGATG 2909
 69 GTTTCGCAACGTAAGACCTTTCGATGATGATGATGATGATGATGATGATGATGATGATG 23

RESULT 2
 US-10-282-122A-40936
 Sequence 40936, Application US/10282122A
 Publication No. US20040029129A1

GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zykard, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: EUTRA 034A US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 40936
 LENGTH: 852
 TYPE: DNA
 ORGANISM: *Vibrio cholerae*
 US-10-282-122A-40936

Query Match 29.0%; Score 842.4; DB 13; Length 852;
 Best Local Similarity 99.3%; Pred. No. 1,26-267;

FEATURE:	NAME/KEY: misc.feature	
LOCATION: (51502)..(51602)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (51786)..(51786)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (51805)..(51805)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (55369)..(55369)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (65309)..(65309)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (65313)..(65313)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (80024)..(80024)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (100091)..(100091)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (102696)..(102696)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (105121)..(105121)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (107248)..(107248)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (117136)..(117136)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (119750)..(119750)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (119924)..(119924)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (120038)..(120038)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (121344)..(121344)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (122167)..(122167)		
OTHER INFORMATION: n equals a, t, g or c		
FEATURE:	NAME/KEY: misc.feature	
LOCATION: (122336)..(122336)		
OTHER INFORMATION: n equals a, t, g or c		

QY	236	GTGTGCACTGGCGGCAATGCTCTTCCAGCGCGGCTTATGGGCTGTGATCACCGCCCAATGTC	355
QY	236	GATTAATGATCTTCTTACAAATTTGATCTGTTCCTTGCTGACCCCTTTATTTATTCAAA	295
QY	236	GCGCATGATATTTTCTATATATGATCATTTCTTACAAAGAACCACTTATTTATTCGCG	355
DB	958326		958385
QY	176	GAGCAAGCTCTGACTTTGTATTTGCGCGGCTTTTAAAGTGTAGTATCGGTGCGAGTT	235
DB	958266	GATCAAGTGTATAGCTTACTTTTCAACGGTTTATGCGGGGTTTTATTTAGCGGACGTGT	958325
QY	116	CTTTTGTGATGTGGTTGGCCAAATGCGGAGCGGATGCGCGGCAATGTTGAACGCT	175
DB	958206	GTTTTTCACGCTGGCTTGGCGGTGCGCGTGTATCCGCCCAATAGCGGTTGACAGTA	958265
QY	62	TCGATCGGC-----CCCTTAGCGGTGCGGTGATGCTGATATTTGGTGGTTTC	115
DB	958146	ACGCTTGGCGATAGTATATATGCGCTTACGTTGATGCTTATGCTTATACCTTTAGGTTT	958205
QY	2	AGAGCTTTGTTATGCTCAGGGTATCTGCAATTTCCCAATATGACCCCGTATGTTT	61
DB	958086	ATAAGGAATAATATGATTAATTCATTAATTTACTTCTCCCACTTGTGATCCGATATTTT	958145
QY	25.6%;	Score 744.2; DB 15; Length 1830121;	
	Best Local Similarity	66.7%; Pred. No. 1.4e-23;	
	Matches 1134; Conservative	0; Mismatches 538; Indels 27; Gaps	4;

Db 958386 GTTGGAGAGTGGATGTGCTTCCACGCTGAGCTTAATGATGTAATTTGGCATGAT 958445
QY 356 TGGATATGGCGTAAACCAACGACCTTCTTTGGTGTGGCCATTTTGTGGCCCTTAA 415
Db 958446 TGGACATCTTAATCTCAAAACGTAATTTTGGCAACGGCTAATTTTGTGGCCCTTAA 958505
QY 416 GTGCAATTCGGTTTGGGATGGGACGTATCGTAATCTTAATGTAATGTAATTTGGGA 475
Db 958506 ATTCCGTTGGTTAGGTTTAAAGGATGTGTAATTTCTTAATCTTGAACATATGGGA 958565
QY 476 CGAGTAACGATGTGCTTGGGCTTTGTAATTCCTTAATGTGGCC--CACTGGCCGC 532
Db 958566 CGGCAACGAAATGCTTGGGCAATATTTTCCGAAATATCTCTTTTACGCTTCGT 958625
QY 533 CATCCTTACAGCTTTTGAATTCGCTTGAAGGCGGTGCTGCTTATCTTAAT 592
Db 958626 CATCATACACACTTTTGAAGCTTTTGAAGGCTGCTGCTTATTCAGATTCGAAT 958685
QY 593 TGGTTATGCTAAACCTTCCTCGCTAGGACGCTATCCGACTTTTATAGCTGATAC 652
Db 958686 ATTTTATTAATAAACCAACGCTCAATGCTTCTGCTGAGTTTATCTTAATGCTTAT 958745
QY 653 GGTACATTCGCTTCCTTGGGATTAAGTCCGTAAGGCAATGCTCACTGGGCTGTT 712
Db 958746 GGGCTCTTCGTTTATTTTGGAAATATGTGCTGAACCTGAAG--TTGAAATTTTC 958799
QY 713 GGTGGCTTCAATTCATAGGCGCAATCTCTCTTACCTTATGATCATCGTATTTTG 772
Db 958800 TTTGGATTAATTAACGAGGCAAGCCCTTGTCTGCGCATGATATATGCTGCTTC 958859
QY 773 ATGATGTTGGTCTTACAAAGCGGCTTGTATCAAGCCGTAGCAGCAAAATGGGT 832
Db 958860 ATTATGCTTGGCTTATTCACGC-----AAAGTCCGCTAATAATGAG 958907
QY 833 AGTAGGTGAACAGTATTAATGATCTTGTACGAGCATCGTCATCAAGTGTGGGT 892
Db 958908 ATTTTATGAAGCAATCTTGAAGCTTTGCGCGCATTTTATGTAAGGAGGATGGGT 958967
QY 893 GAAATGACGAAACGCGCAAGCGTGTGTAATGATGTAATGATGATGATGATGAT 952
Db 958968 GCTATGACGCTACAGGTAAGCATTTGCTCAAGCTCATTAATGATGATGATGAT 959027
QY 953 GTGGGCAACATCGTTCCTCTAGTACTACACGAGGTTTGGAAAGTGGCGGA 1012
Db 959028 GTGGCAATATCAATTTCCGCTGATTAATCCGTAAGTATGGAAGGCGGAT 959087
QY 1013 GCCAGTTGCTGGCTATATTTCTGTAGATATGCGGAGTATTTGCCAATTAGT 1072
Db 959088 GCTGAATTTTAAAGTTATATTCGTGATATGACATGCCGTGATTTCCGCGCACTGGC 959147
QY 1073 ACCAAACCTGGGATGCTAATGCCAATTTAAACCAAGATGCTCAACAAATCTTACCGT 1132
Db 959148 ACGAAACCTGGGATGCTAATGCCAATTTAAACCAAGATGCTCAACAAATCTTACCGT 959207
QY 1133 AAAGTAGAGTACATGAGGACGCGTATGATGATGATGATGATGATGATGATGAT 1192
Db 959208 AGAGGCTGATGATATGCGGCGGCTATGCTGTAAGGCAAGAGATGCGCTAACCT 959267
QY 1193 GATGCTGCTATATGACCACTGTAAGAAAGATGCTGATGATGATGATGATGAT 1252
Db 959268 AATGGAACAACTATGATGATGATGATGATGATGATGATGATGATGATGAT 959327
QY 1253 GACGAGGTGAATTTCTTAATCTTACATCCGCGGTAATTTCAATGCGGCTTGGCG 1312
Db 959328 GATGAGAGAGATTTTAATCTTTTAACTTGGGAAATTTGATCTTGGTGTCTTCGT 959387
QY 1313 CTTTGCATGTAACGCACTATTTTCAATGCTGGGAGATACCTTGAATCTCAACAGTACT 1372
Db 959388 CCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959447
QY 1373 CAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1432
Db 959448 CAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959507

QY 1433 CTGGGCTGATGGCAGATCAAGGAAAAAGCCGGCTTGGCGTATCAAGATGCTC 1492
Db 959508 TTAGCACTTATGCTCAGATCAAGGCAAAAAAGCCGCAAGCATATCAATTAATTTTG 959567
QY 1493 AATGCGACATTTTACCAAGATCACTGCAATGATGCGGATGTCAGCTTAAACGTGAG 1552
Db 959568 AATGCGACATTTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 959627
QY 1553 CCATTCACAGGCGCTCAGTTCATATCAATCCAAAGATTTAAACGTGAGATTTGGA 1612
Db 959628 CCTTCCATTAACCAAAATTAAGAAATTAATCAAGTATTAAGAAACCTTGAAGATTGGA 959687
QY 1613 ACTTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1672
Db 959688 ACTTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959747
QY 1673 TACCGTTTCACTGATTAAT 1691
Db 959748 TATCCTTTTCCGCTAAT 959766
RESULT 4
US-10-329-670-1
; Sequence 1, Application US/10329670
; Publication No. US2004018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Hemophilus influenzae Rd Genome, Fra
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; PRIOR FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

```

```

NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature

```



```

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 30692
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-10-282-122A-30692

Query Match      17.4%; Score 505.4; DB 13; Length 852;
Best Local Similarity 74.6%; Freq. No. 7.9e-156;
Matches 655; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY      840 TGAACAGATTTAGATCTTGTGACGCGCATGTCATGATCAAGTGTGTGGTTGAATG 899
DB      2 TGAACAGATTTAGATCTTGTGACGCGCATGTCATGATGAGAGTGTGTGAATG 61

QY      900 AACGAAACGGGCAAGCTGTGTGACGCGCATGTCATGATGAGAGTGTGTGAATG 959
DB      62 CACGGAATGCGCAAGGTTGTGACGCGCATGTCATGATGAGAGTGTGTGAATG 121

QY      960 AACATCAGTTCTCTAGTGACTACACGCAAGATTTTGGAAAGTGCCTGAGCGAGT 1019
DB      122 ATAAACCAATTCGCTGATGACACACGCAAAAGCTATGGAAGCGGCAATTCAGAT 181

QY      1020 TGCCTGCTATATTTGCTGTGATGATGATGATGATGATGATGATGATGATGATG 1079
DB      182 TTTTGGTTATATCCGTGCTATGATGATGATGATGATGATGATGATGATGATGATG 241

QY      1080 CCTGGGATGCTATGCTCAATTTAAACCAAGCATGCTCAACATCTTACCTTAAGGTG 1139
DB      242 CCTGGGATGCTCAATGCTATGAAACAGTGCATGCTCAATTAATCCGCAACCGCAAGCA 301

QY      1140 AAGATGACATGCGGAGCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1199
DB      302 CCGATGATGCGGAGCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 361

QY      1200 GTCATATTCACCATGTAAGAAAGATGTTGATGATTTGACCGGTGCGTTGATGACCGAG 1259
DB      362 AAACGCTGATCAATTCGCAAAATCGTCAATATCTGCCAACGCGATGATGACCGAG 421

QY      1260 GCGAATTTCTTAACCTTCAACATCCGGGTGATTTCAACAGGGGTGTTGGCGCTTGA 1319
DB      422 GCGAATTTCTTAACCTTCTCAACCCCGGTGATTTGAACGCGGTGCTTACGCTTGA 481

QY      1330 TGTACAGCATCATTTTTCATTTGCTGCGGGAATACCTTGTATCTCAACAGTACTACGCTT 1379
DB      482 TGTACAGCATCATCTTCTTCAACCCCGGTGATTTGAACGCGGTGCTTACGCTTGA 541

QY      1380 CATGATGCTGCTGCGGTGATTTCAACAGGTCGAGTTTATGTTGTTCTTCTGCGC 1439
DB      CATGATGCTGCTGCGGTGATTTCAACAGGTCGAGTTTATGTTGTTCTTCTGCGC 1439
```

```

DB      542 CTCTGTATGTCGCGCTTGCTGATTTCAACCAATTTAAGTGTTCACCTTTCTTGCT 601
QY      1440 TGATGCAACAGATCAACAGGAAAAAGCCGCGCTTGCGTATCAGAGATGCTGAATGCGC 1499
DB      602 TGATGCAACAAATTAACGCGCAAAAAAGCGGTAAAGCGTACCACAAAATCAATGACAC 661
QY      1500 ACATTTCAAGATCAACCTGGAATTGATGCGGATGTCAGGTAAAGTGGACCATTC 1559
DB      662 ACATTTTGAAGATCAACCTGATCTCATGAAAAATGTGCAATTTAAAGCTGACCCCTTC 721
QY      1560 CAGCGCTCAAGTTCCATATCATCAATCAAGATTTAAACACTGCAAGATTTGAAACTTGG 1619
DB      722 CGTGCCACAAATTAACATTAACCCCGAGATTTAAACTCTCCAAAGATTTAGAACTTGG 781
QY      1620 TCATTTGATGATTTTGTGACGTCAGCGATATCAGTTCAGCATCTTCAATTCATCCCT 1679
DB      782 TGACTATGATGATTTTAAAGTACGCGCTATCATGCGCATGCTCAATTAATTCCTT 841
QY      1680 TTTCACTTAA 1690
DB      842 TCTCTGTGTAA 852

RESULT 6
US-10-282-122A-22162
; Sequence 22162, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OR INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCES: EPI-TRA-034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 22162
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-10-282-122A-22162
```


QY 1260 GTGAATTTCTTACTTCTTCAATCCGGGTGAATTTACATGGGGTGTGGCCCTTCCA 1319
DB 422 GTGAGATGACTCTTTTATATATCCCGAGAAACAGCGTTAGGGTGTTCAGTCCATGCA 481
QY 1320 TGTACAGCATATTTTCAATGCTGGGGATACCTTGTATCTCAACAGTACTCAGCGTT 1379
DB 482 TGCATACATATCTTTTTCATGCTGGGTGACAGGCTTATTTAACCTCATACAGTA 541
QY 1380 CATGTGATGTCCTTGGGGTGTGAATTTCAACATGCTGAGGCTTATGTTCTTGGCC 1439
DB 542 GGTGTGATGTCCTTGGGGTGTGAATTTCAACATGCTTGTATTTATTTACCGC 601
QY 1440 TGTATGCAAGATCAAGAGAAAGAGCGGCTTGGCGTATCAAGATGCTCAATGCGC 1499
DB 602 TGTGTGCTCAATCAAGAGGATTAACAGGTAAGATTCATTAAGATGTTATGCTC 661
QY 1500 AATTTTACCAAGATCACTCGAATGATGCGCGATGTCAGCTTAAACGTGAGCCATTC 1559
DB 662 AATTTTATGAGAACCAATTAACCTGATGCTGATGTTCACTTAAAGAGAGCCATTCG 721
QY 1560 CAGCGCTCAGTTCATATCAATCCAAAGATTTAAACAGTCAAGATTTGGAACTTGG 1619
DB 722 CATTCCTAGCTATATATTAATCCAAATTAAGATGATGATTAAGAACTTGG 781
QY 1620 TCACTTTGATATTTTGAAGCTCAGCGATATCACTTCAAGATCTTATCAATACCGT 1679
DB 782 TCAGCAGATGATCTTGTAGTGTAAAGTTATGATGCTCAGAGCGATTAATACCGT 841
QY 1680 TTTTCACT 1686
DB 842 TCAGCT 848

RESULT 8
US-09-741-669-182
; Sequence 182, Application US/09741669
; Patent No. US200202278A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Gene identified as required for
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 182
; LENGTH: 876
; TYPE: DNA
; ORGANISM: *Bacterichia coli*
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(876)
US-09-741-669-182

Query Match 8.5%; Score 248.2; DB 9; Length 876;
Best Local Similarity 58.8%; Pred. No. 1.7e-70;
Matches 510; Conservative 0; Mismatches 288; Indels 69; Gaps 2;

QY 14 ATGCTCAGGTTATCTGCAATTTCCCAATATTAACCCGATTTGTTTGAATGCGCCT 73
DB 1 ATGACCACTACTCTATCTGCAATTTCCGAGTTGATCCGATTTCTCAATTAAGACCC 60
QY 74 CTAGCGTGCCTGATGCTTATGCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 133
DB 61 GTGGGCTTCACTGATGAGGCTGATGATGATGATGATGATGATGATGATGATGATG 120
QY 134 GCATATCCGAGGCGATGCGGCGGAGTGTGAGCGGTGAGCAGTCTGTGACTTG 193

DB 121 GCAACACAGCGGGAATGCTCCGGCGAGCGGCTGAGCCAAATATAGATTGAAACCTTA 180
QY 194 TTTATGCGGCTTTTATGAGTATGATGATGATGATGATGATGATGATGATGATGATG 253
DB 181 CTATATGCGGCTTCTCCGCGCTTCTTCTCCGCGGAGAGATGATGATGATGATGATGATG 240
QY 254 AATTTTATGCTTCTGCTGAGCCGCTTATTTTAAAGTGTGAGCTGGCGGATG 313
DB 241 AATTTTCCGAGTTTATGCGGAGTCCGCTTATCTGCTGCTGCTGAGCGGAGCTG 300
QY 314 TCTTTCACAGCGGCTTATGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 373
DB 301 TCTTTCACAGCGGCTTATGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 360
QY 374 CAAGCAAGCTTCTTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 433
DB 361 AAGCTTCTTCTTCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 420
QY 434 ATGGAAGTATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
DB 421 GCGGCGGCTGAGGCACTTATTAAGGATGATGATGATGATGATGATGATGATGATGATG 480
QY 481 -----AAGGAT 487
DB 481 CCGTTGCGATGCTTCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 540
QY 488 GTGCTTGGCTTTTGTATTCCTTAATGATGATGATGATGATGATGATGATGATGATGATG 547
DB 541 CCGGAGTGGCAATTCATTTTCACTTATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 548 TATGATGCTTATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 607
DB 601 TACAGCTGCTGCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 608 CCGTGTGCTGAGGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 667
DB 661 CCAAGCCCAATGAGGAGCTGCTCAGGTTGTTCTTATGATGATGATGATGATGATGATGATG 720
QY 668 CTGTGGAATACGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 724
DB 721 ATGTGATGATTTTCCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 780
QY 725 TCAATGAGGCAATCTCTCTTACCTATGATGATGATGATGATGATGATGATGATGATGATG 784
DB 761 AGCATGAGGCAATCTCTCTTACCTATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 765 TCTTACAGAGCGGCTTGTATCAAGAC 811
DB 841 GCATATGCTGAGCCACAGCAACAC 867

RESULT 9
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: *Buchnera sp.*

US-09-790-988-1

Query Match 5.2%; Score 152.2; DB 9; Length 640681;
 Best Local Similarity 52.6%; Pred. No. 1.7e-16;
 Matches 432; Conservative 0; Mismatches 323; Indels 66; Gaps 2;

QY 35 TTTCCTCAATATGACCCCGTATGTTTGCATGCGCCCTTACCGGTGCGGTATGCG 94
 DB 479304 TTTCCTCAATATGACCCCGTATGTTTGCATGCGCCCTTACCGGTGCGGTATGCG 94
 QY 95 TTGATGATTTGGGGGTTTCCCTTTTGCATGCGGTGCGCATGCGCGAGATCCG 154
 DB 479364 TTGATGATTTGGGGGTTTCCCTTTTGCATGCGGTGCGCATGCGCGAGATCCG 154
 QY 155 GCGGGGAGTGTGAGCGCGTGAAGTCTGACTGTTGTTTCCGCGCTTTTAAAGT 214
 DB 479424 AATGAAAAATATGATGATGATGAAAAAATGAAAAATATATATATATATATAT 479483
 QY 215 GTATGATCGGTGCGGAGTGTGATGATCTTCTACATTTTGTATCTGTTCTGCT 274
 DB 479484 TCAATGATCGGAGGAAAGATGATATATATATATATATATATATATATATAT 479543
 QY 275 GACCTCTTATTTATTTCAAGTGTGACTGCGGAGTCTTCCAGCGCGCTTATG 334
 DB 479544 AATATGATCTGATTTTATATATATGGAAGAGCATGCTATTTATGAGATTTA 479603
 QY 335 GGTGTATACCGCCATGTTCTGATGCGCGTAAACCAACGACCTTCTTGTGTG 394
 DB 479604 GGTGTATATCTGATGCTGATGCTTCTTAAATATGAAAAAATATATGAAATA 479663
 QY 395 GCGGATTTGTTGCGCCCTTATGAGCATGCGGTGCGGATGCGGATGCTTACTT 454
 DB 479664 TCGATTTATCTCTCACTCACTATACCTTTGTTGAGCGCTGATGAGAACTT 479723
 QY 455 ATGATATGTAACCTTTGGGACAGT---AAGGATGCTGCTTGGCTTTGATTCCT 511
 DB 479724 ATTAATATGTAACCTTTGGGACAGT---AAGGATGCTGCTTGGCTTTGATTCCT 479783
 QY 512 AATGCT----- 517
 DB 479784 AATTCCTCAATATGATTTAAAGAAATTAACCGAATACCACTATATTA 479843
 QY 518 -----GGCCACTGCGCGCCATCCTTCAAGCTTTATGAAATGCGCTTAGAGGC 568
 DB 479844 GATCAATAGGAGCATTTGCGGCTATCTCAACCATTTATGCAATTTTATGAGG 479903
 QY 569 GTGCTTCTGTTTATTTCTTATTTGTTTATGTTGTTGTTGTTGTTGTTGTTG 628
 DB 479904 ATTCCTCTGTTTATTTATTTATTTTTCGAAAAAAGATAGACCAAGCTAGTAT 479963
 QY 629 TCCGACTGTTTATGCTGATACGATCCGCTTCTGCTGAGTATGCTGCTGAG 688
 DB 479964 ACGGTTTATTTTATTTTATTTATGATTTATGAAATTTATGAAATTTTATGAA 480023
 QY 689 CCAATGCTCACTGCTGCTGCTGCTGCTTCAATGAGGCAATCCCTCTCTTA 748
 DB 480024 CCAATGCTCACTGCTGCTGCTGCTGCTTCAATGAGGCAATATATATCACTT 480083
 QY 749 CCTATGCTGATCATGCTGCTTGTGATGATGCTTGTGCTTA 789
 DB 480084 CCTATGCTGATCATGCTGCTTGTGATGATGCTTGTGCTTA 480124

RESULT 10
 US-10-672-787-35
 ; Sequence 35, Application US/10672787
 ; Publication No. US20040067554A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LAGACE, Robert, E.
 ; APPLICANT: PATTERSON, Chandra
 ; APPLICANT: BERG, Kim, L.
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
 ; FILE REFERENCE: ELITRA.025C1

CURRENT APPLICATION NUMBER: US/10/672,787
 ; CURRENT FILING DATE: 2003-09-26
 ; PRIOR APPLICATION NUMBER: 09/596,002
 ; PRIOR FILING DATE: 2000-06-16
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 35
 ; LENGTH: 96109
 ; TYPE: DNA
 ; ORGANISM: Moraxella catarrhalis
 US-10-672-787-35

Query Match 4.9%; Score 142.2; DB 13; Length 96109;
 Best Local Similarity 59.1%; Pred. No. 8.5e-34;
 Matches 262; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

QY 38 CCAATATGACCCCGTATGTTTGCATGCGCCCTTACCGGTGCGGTATGCGCTTG 97
 DB 51130 CCAAGATTAACCTGTTGCTTTAGACTGGGCAATGATGATGATGATGATGATG 51189
 QY 98 ATGATTTGTTGCTTCTTTGCTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 157
 DB 51190 ATGATTTGTTGCTTCTTTGCTATGTTGCTGCTGCTGCTGCTGCTGCTGCTG 51246
 QY 158 GCGAGTGTGAGCGCGTGAAGCATCTGACTGTTATTCGCGGCTTTTATGATGTA 217
 DB 51247 GCAACAGATTTTACCTCGATGATGTTGCTGATGTTGTTTGTGCTGATGCTG 51306
 QY 218 GTATGCTGCGCGAGTGTGATGATCTCTACATTTTATGATCTCTGCTGCTGCTG 277
 DB 51307 ATTTAGTGTGCTGATGCTGATGCTTATTAACAATTTGTTGATGCTGATGCTG 51366
 QY 278 CCTTTATTTATTTCAAGTGTGACTGCGGCAATGCTTCCAGCGGCTTATGCTG 337
 DB 51367 CCTTTATTTATTTTCAAGTGTGACTGCGGCAATGCTTCCAGCGGCTTATGCTG 51426
 QY 338 GTATGACCGGCAATGCTGCTGATGCGGCTTAAACCAACGACCTTCTTGTGTG 397
 DB 51427 GTGTTGCTTGCAGTGTGATTTTGCACATTAATTAACCAACGACCTTCTTGTG 51486
 QY 398 GATTTGTTGCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 457
 DB 51487 GATTTGTTGCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 51546
 QY 458 AATATGTAACCTTTGGGACAGT 480
 DB 51547 AATGCGAGCTTTGGGACAGAT 51569

RESULT 11
 US-10-297-465A-1/c
 ; Sequence 1, Application US/10297465A
 ; Publication No. US20040142413A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Simpson, Andrew
 ; APPLICANT: Reinach, Fernando
 ; APPLICANT: Serubal, Joao
 ; APPLICANT: Medians, Joao
 ; TITLE OF INVENTION: Isolated Genome of Xylella fastidiosa and Uses Thereof
 ; FILE REFERENCE: PABESP 202 US (10213376)
 ; CURRENT APPLICATION NUMBER: US/10/297,465A
 ; CURRENT FILING DATE: 2001-06-07
 ; PRIOR APPLICATION NUMBER: PCT/IB01/01618
 ; PRIOR FILING DATE: 2001-06-07
 ; PRIOR APPLICATION NUMBER: 60/209,906
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1
 ; LENGTH: 2731748
 ; TYPE: DNA
 ; ORGANISM: Xylella fastidiosa

PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: US 60/330,462
PRIOR FILING DATE: 2001-10-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2221
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1988
LENGTH: 3950
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURES:
OTHER INFORMATION: Genbank Accession No. NM_053380
US-10-152-319A-1988

Query Match 2.7%; Score 77.8; DB 12; Length 3950;
Best Local Similarity 56.7%; Pred. No. 1.9e-13;
Matches 164; Conservative 0; Mismatches 122; Indels 3; Gaps 1;

QY 2480 TCTAGAGTGAAGATCAGACCGGAGAAAGATTGAAAAATCGTGATGGGCACTG 2539
DB 724 TCCGAGGTAAAGACCAACACAGAGCCAGTTGAAGATCATGACAGTAGCTCCT 665
QY 2540 GCGAAGCAGCGGAAACTCTTCTTTACAGGCATATGCCAGCTGAGAGATATG 2539
DB 664 GCAATATCCCTTCTGAATCTATTTCTCTCTGCTGCTGCAATAGTCCCAATGTTG 605
QY 2600 GTCACTAGTACCAATATTTGGCACCATCACCATTAGAAATCGCGTTTCAAC---CGGT 2656
DB 604 GTGATGAGGTCCCGATGTTAGACCCCATATGATGGGATGAGCAGCGGAGCTGACAG 545
QY 2657 AACCCACCGGCAAGACCAACATATATGAAGTACCGCTTGAAGATTGAATCAGT 2716
DB 544 AAGAGAGGCAACCATGTCAGATGATTGACAGAGACGTGTGAGCTTTGACCACAG 485
QY 2717 GCCGTGCCACTAAACCAATCATCATCTGCAATTTGGGGAGAGCA 2765
DB 484 ACTGTCAACGACACCCCTATCAAGATCCAGCCACAGATTGACATTA 436

RESULT 14
US-10-276-774-1331/c
Sequence 1331, Application US/10276774
Publication No. US20040053245A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
APPLICANT: Ieng, Y, Tom et al
TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-030
CURRENT APPLICATION NUMBER: US/10/276,774
PRIOR FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 2700
SOFTWARE: Custom
SEQ ID NO 1331
LENGTH: 2208
TYPE: DNA
ORGANISM: Homo sapiens
US-10-276-774-1331

Query Match 2.6%; Score 75.4; DB 13; Length 2208;
Best Local Similarity 56.0%; Pred. No. 7.6e-13;
Matches 164; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 2476 CATTTCTAGAGTGAAGATCAGACCGGAGAAATTTGAAAAATCGTGATGGTGC 2535
DB 738 CACCTCCAGGGGCAAGACCAACAGGACAGCCAGTTGAAGATCATGACAGTGC 679
QY 2536 ACTGGGAAAGCAACGCGAAACTCTTTTACAGCCCATATGGCCAGGCTACAGAGT 2595
DB 678 TCCGTCAAAAGCTCTTCTGAACTCACTCCGATCTCCACCTGCATGAGGCAACATAGT 619

QY 2596 ATTGCTCAGATGATACCAATATTGGACCCATCACCATAGAGATCGGGTTCAACCG- 2654
DB 618 GTTGGATGATGACGTTCCATGTTGGCCCCCATGTATATGGGAGTGGACCCGGAACGT 559
QY 2655 --GTAAACCCAGGCAAGACGACCAACATATATGAAGTACACGTTGATGAGATTGAAT 2712
DB 558 GAGCAATGAAAGAGACACCATCTGACACATGACATGACGTTGAGGTGCTGAGACTCTGCAC 499
QY 2713 CAGTCCGTTGCCACTTAAACCAATCATCATCTTGCATTTGGGTGGAGACAA 2765
DB 498 CAAGACGTCACACACACCCCGATCATCAGCCCAACAAAGGTTGACATTA 446

RESULT 15
US-10-173-999-149/c
Sequence 149, Application US/10173999
Publication No. US2004005563A1
GENERAL INFORMATION:
APPLICANT: Mack, David H.
APPLICANT: Gish, Kurt C.
APPLICANT: Bos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
FILE REFERENCE: 018501-002420US
CURRENT APPLICATION NUMBER: US/10/173,999
PRIOR FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: US 60/299,234
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: US 60/315,287
PRIOR FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/372,246
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 163
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 149
LENGTH: 2280
TYPE: DNA
ORGANISM: Homo sapiens
US-10-173-999-149

Query Match 2.6%; Score 75.4; DB 16; Length 2280;
Best Local Similarity 56.0%; Pred. No. 7.7e-13;
Matches 164; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 2476 CATTTCTAGAGTGAAGATCAGACCGGAGAAATTTGAAAAATCGTGATGGTGC 2535
DB 771 CACCTCCAGGGGCAAGACCAACAGGACAGCCAGTTGAAGATCATGACAGTGC 712
QY 2536 ACTGGGAAAGCAACGCGAAACTCTTCTTTACAGGCATATGGCCAGGCTGACGAGT 2595
DB 721 TCCGAAAAGCTTCTTGAATCACTCCATTCCTCCACCTGACATGAGCGCAATAGT 652
QY 2596 ATTGCTCAGATGATCAATATTTGGCACCATCATCAATAGAAATCGGGTTCAACCG- 2654
DB 651 GTTGGATTGACCTTCCATATGTTGGCCCCCATATATGAGGATGACGCCCAACAGT 592
QY 2655 --GTAAACCCAGGCAAGACGACCAACATATATGAAGTACACGTTGATGAGATTGAAT 2712
DB 591 GAGCAATGAAAGAGACACCATCTGACACATGACATGAGTGAAGTCTGAGCTCTGCAC 532
QY 2713 CAGTCCGTTGCCACTTAAACCAATCATCATCTTGCATTTGGGTGGAGACAA 2765
DB 531 CAAGACGTCACACACCCCGATCATCAGCCCAACAAAGGTTGACATTA 479

Search completed: July 31, 2004, 19:10:09
Job time : 903 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 07:32:25 ; Search time 761 Seconds
(without alignments)
16239.170 Million cell updates/sec

Title: US-09-700-712A-1

Perfect score: 2909

Sequence: 1 gaggaaggttctatgcctc.....cagcgaatcgcgctcgaag 2909

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002s:*
- 7: geneseqn2003as:*
- 8: geneseqn2003bs:*
- 9: geneseqn2003cs:*
- 10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2909	100.0	2909	AAZ40646	AAZ40646 Nucleotide
2	1222	42.0	1222	AAZ40648	AAZ40648 Nucleotide
3	1112.6	38.2	1149	ACFA53067	ACFA53067 Prokaryot
4	842.4	29.0	852	AAZ24801	AAZ24801 Vibrio ch
5	842.4	29.0	852	ACFA53066	ACFA53066 Prokaryot
6	838	28.8	838	AAZ40647	AAZ40647 Nucleotide
7	744.2	25.6	7	AAZ42063_09	AAZ42063_09 Nucleotide
8	505.4	17.4	852	ACFA28222	ACFA28222 Prokaryot
9	462.2	15.9	849	ACFA44455	ACFA44455 Prokaryot
10	403.8	13.9	849	ACFA44455	ACFA44455 Prokaryot
11	292.2	10.0	807	AAZ37118	AAZ37118 Nucleotide
12	278.2	9.6	110000	ACF67867_27	ACF67867_27 Nucleotide
13	278.2	9.6	110000	ACF65386_10	ACF65386_10 Nucleotide
14	278.2	9.6	110000	ACF65386_1	ACF65386_1 Nucleotide
15	277.8	9.5	876	ACF69784	ACF69784 Phototroph
16	248.2	8.5	876	AAH81383	AAH81383 Escherich
17	208.6	7.2	834	AD332523	AD332523 DNA encod
18	208.4	7.2	834	ABG90157	ABG90157 M. capsul
19	161.2	5.5	852	AAZ53292	AAZ53292 Neisseria
20	156.4	5.4	852	AAZ53291	AAZ53291 Neisseria
21	156.4	5.4	102634	AAH81464	AAH81464 N. mening
22	156.4	5.4	110000	AAH81490_10	AAH81490_10 Nucleotide
23	156.4	5.4	349980	AAZ21609	AAZ21609 Neisseria

24	152.2	5.2	110000	6	ABA92787_4	Continuation (5 of
25	149.2	5.1	849	7	ABZ40310	ABZ40310 N. gonorr
26	149.2	5.1	852	3	AAZ53290	AAZ53290 Neisseria
27	148.6	5.1	1515	4	AAAF61079	AAAF61079 P. putida
28	145.4	5.0	420	6	ABN26154	ABN26154 Human ORF
29	142.2	4.9	96109	4	AAZ28548	AAZ28548 Genomic I
30	137	4.7	420	6	ABN26172	ABN26172 Human ORF
31	84	2.9	2440	9	ABT41887	ABT41887 Toxicity
32	84	2.9	2440	9	ABD58174	ABD58174 Toxicity
33	84	2.9	2440	9	ABD52673	ABD52673 Primary r
34	83.6	2.9	4473	5	AAH88385	AAH88385 DNA encod
35	79	2.7	2075	9	ABD58991	ABD58991 Toxicity
36	79	2.7	2075	9	ABD53745	ABD53745 Primary r
37	77.8	2.7	3950	7	ABT42286	ABT42286 Toxicity
38	75.4	2.6	2208	4	ABA09555	ABA09555 Human Na-
39	75.4	2.6	2208	4	AAK52748	AAK52748 Human pol
40	75.4	2.6	2280	9	ADB80585	ADB80585 Ovarian c
41	75.4	2.6	2288	2	AAV5498	AAV5498 Human sod
42	75.4	2.6	2380	4	AAK51764	AAK51764 Human pol
43	75.4	2.6	2520	4	AAK94823	AAK94823 Human ful
44	75.4	2.6	3952	8	ACH03831	ACH03831 Human CDN
45	75.4	2.6	4137	3	AAA75484	AAA75484 DNA encod

ALIGNMENTS

RESULT 1
ID AAZ40646 standard; DNA; 2909 BP.
XX
AC AAZ40646;
XX
DT 08-MAR-2000 (first entry)
XX
DE Nucleotide sequence of V. cholerae *thyA* gene.
XX
KW Vibrio cholerae; *thyA*-negative strain; Deltathya; *thyA* gene; vaccine;
KM cholera; antimicrobial therapy; ss.
XX
OS Vibrio cholerae.
XX
PN WO9961634-A1.
XX
PD 02-DEC-1999.
XX
PF 21-MAY-1999; 99WO-EP003509.
XX
PR 26-MAY-1998; 98SE-00001852.
XX
PA (SBLV-) SBL VACCIN AB.
XX
PI Carlin N, Lebens MR;
XX WPI, 2000-062719/05.
XX P-PSDB; AAY59126.
XX
XX New Vibrio cholerae strain defective in the *thyA* gene, for use in
PT vaccines and for recombinant protein production.
XX
PS Claim 9, Fig 1, 42pp; English.
XX
XX The invention provides a method for producing a Vibrio cholerae *thyA*-
CC negative strain which is Deltathya strain lacking *thyA* gene functions.
CC The method comprises site-directed mutagenesis of the V. cholerae
CC chromosome to delete and/or insert nucleotides at the *thyA* locus; The V.
CC chromosome *thyA*-negative strains are used: for overproduction of
CC recombinant proteins; and in vaccines to prevent or treat cholera (or
CC other diseases, if engineered to express the appropriate proteins). The
CC *thyA* gene is also useful for insertion of foreign genes, in a selective
CC and site-specific manner, and the proteins expressed by the *thyA* gene or
CC by its 5'-flanking region, are useful in research and as targets for
CC antimicrobial therapy. When used for recombinant protein production, V.

CC cholerae provides high yields with secretion of products into the culture medium for ease of subsequent recovery. The thya-negative strain can be maintained by thymine complementation, eliminating the need for antibiotic selection. The present sequence represents the nucleotide sequence of the V. cholerae thya gene

XX

SO Sequence 2909 BP; 734 A; 644 C; 737 G; 794 T; 0 U; 0 Other;

Query Match 100.0%; Score 2909; DB 3; Length 2909;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GAGAGGTTGTTATGCTCAGGAGTATCTGACATTTCCCAATATTGACCCCGATTGTT 60
DB 1 GAGAGGTTGTTATGCTCAGGAGTATCTGACATTTCCCAATATTGACCCCGATTGTT 60
QY 61 TTGATCGGCGCTTACGCGGTGCGCTGTATGCTTATGTTTGGTGGTTTCTTTT 120
DB 61 TTGATCGGCGCTTACGCGGTGCGCTGTATGCTTATGTTTGGTGGTTTCTTTT 120
QY 121 TGTATGTTGTTGGCCAAATCGCGAGCGGATCGCGGCGAGTGTGTAAGCGGTGACA 180
DB 121 TGTATGTTGTTGGCCAAATCGCGAGCGGATCGCGGCGAGTGTGTAAGCGGTGACA 180
QY 181 AGTCTCTGACTTGTATTTGCGCGCTTTTATGCTGTATGATCGGTGCGCAATTGTTA 240
DB 181 AGTCTCTGACTTGTATTTGCGCGCTTTTATGCTGTATGATCGGTGCGCAATTGTTA 240
QY 241 TGTGATCTTCAATATTGATCTGTTCTGTGCTACCTCTTATTTATTAAGATGTG 300
DB 241 TGTGATCTTCAATATTGATCTGTTCTGTGCTACCTCTTATTTATTAAGATGTG 300
QY 301 GACTGCGGCGAGTCCCTTCCAGCGCGCTTATTTGGGTGTGATCACCGCGCATTTCTGTA 360
DB 301 GACTGCGGCGAGTCCCTTCCAGCGCGCTTATTTGGGTGTGATCACCGCGCATTTCTGTA 360
QY 361 TGGCGCTTAATAAACCAACGACCTCTTCTTGTGTGCGCGATTTGTTGCCCCCTTTAGTGC 420
DB 361 TGGCGCTTAATAAACCAACGACCTCTTCTTGTGTGCGCGATTTGTTGCCCCCTTTAGTGC 420
QY 421 ATTGCGTTGGGATGGAGCTATCGGTAACTTATGATAGTGAATTTGGGAGCGAGT 480
DB 421 ATTGCGTTGGGATGGAGCTATCGGTAACTTATGATAGTGAATTTGGGAGCGAGT 480
QY 481 AACGAGTGTGCTTGGGCTTTTGTATTCCTTAATGATGTCGCCACATGCGCGCATCTTC 540
DB 481 AACGAGTGTGCTTGGGCTTTTGTATTCCTTAATGATGTCGCCACATGCGCGCATCTTC 540
QY 541 ACAGCTTATGATTTGCGCTTAGAGCGGTGCTGTTCTTATTTCTTAATTTGGTTAT 600
DB 541 ACAGCTTATGATTTGCGCTTAGAGCGGTGCTGTTCTTATTTCTTAATTTGGTTAT 600
QY 601 TGGTAAACCTGTCGCTTAGGAGCGTATCGGACTGTTTTTATGCTGATACGTTACAT 660
DB 601 TGGTAAACCTGTCGCTTAGGAGCGTATCGGACTGTTTTTATGCTGATACGTTACAT 660
QY 661 CCGCTTCTTGTGATAGTCCGTAGGACAGATGCTCAGTGGGTGCTTTGGTGGCT 720
DB 661 CCGCTTCTTGTGATAGTCCGTAGGACAGATGCTCAGTGGGTGCTTTGGTGGCT 720
QY 721 CATTTCAATGGGCAATCTCTCTTACTATGATGATCATCGTATTTTGAATGAT 780
DB 721 CATTTCAATGGGCAATCTCTCTTACTATGATGATCATCGTATTTTGAATGAT 780
QY 781 TTGCTCTTACAGCGCGGTTTGTATCAAGCGGTGACAGCAAAATAGGATAGT 840
DB 781 TTGCTCTTACAGCGCGGTTTGTATCAAGCGGTGACAGCAAAATAGGATAGT 840
QY 841 GAAACGATATTAGATCTTGTGAGCGCATCGTCAAGATGCTTGGTTGAAATGA 900
DB 841 GAAACGATATTAGATCTTGTGAGCGCATCGTCAAGATGCTTGGTTGAAATGA 900
QY 901 ACAGACGGGCAAGCGTTTGACTGTATTAATGCCGATTGACCTACGATGTGGCA 960
```

```
DB 901 ACAGACGGGCAAGCGTTTGACTGTATTAATGCCGATTGACCTACGATGTGGCA 960
QY 961 CATATGTTTCCCTTAGTACATACAGCAAGTTTGTGAAAGTGCCTGACCGAGT 1020
DB 961 CATATGTTTCCCTTAGTACATACAGCAAGTTTGTGAAAGTGCCTGACCGAGT 1020
QY 1021 GTCGCGTAAATTCGTGTAGCATATGCGCGGATTTTGGCCAAATTGATACCAAAAC 1080
DB 1021 GTCGCGTAAATTCGTGTAGCATATGCGCGGATTTTGGCCAAATTGATACCAAAAC 1080
QY 1081 CTGGATGCTAATGCCAATTTTAAACCAACATGCTCAACATTCCTTACCGTAAAGTGA 1140
DB 1081 CTGGATGCTAATGCCAATTTTAAACCAACATGCTCAACATTCCTTACCGTAAAGTGA 1140
QY 1141 GGATGACATGGAGCGGTGATGTGTTGATGAGGTGAGGTGAGTGGCTGATGTGG 1200
DB 1141 GGATGACATGGAGCGGTGATGTGTTGATGAGGTGAGGTGAGTGGCTGATGTGG 1200
QY 1201 TCAATATTGACCACTGTAATAAGATTGTGATGATTTGACCGGTGAGTGGCTGATGTGG 1260
DB 1201 TCAATATTGACCACTGTAATAAGATTGTGATGATTTGACCGGTGAGTGGCTGATGTGG 1260
QY 1261 TGAATTTCTTACTTCAATCGGAGTATTTCAATGAGGTGTTTGGCCCTTGCAT 1320
DB 1261 TGAATTTCTTACTTCAATCGGAGTATTTCAATGAGGTGTTTGGCCCTTGCAT 1320
QY 1321 GTCAGGCAATCTTTTCAATGCTGCGGATATCTTGTATCTCAAGATGCTCAAGTTC 1380
DB 1321 GTCAGGCAATCTTTTCAATGCTGCGGATATCTTGTATCTCAAGATGCTCAAGTTC 1380
QY 1381 ATGTGATGTCCTTGGGTTGATTTTCAACATGAGTGTGAGTGTGCTTCTTGGCT 1440
DB 1381 ATGTGATGTCCTTGGGTTGATTTTCAACATGAGTGTGAGTGTGCTTCTTGGCT 1440
QY 1441 GATGCGACATACAGGGAAGGCGGCTTGGCGTATCAACATGCTCAATGCGCA 1500
DB 1441 GATGCGACATACAGGGAAGGCGGCTTGGCGTATCAACATGCTCAATGCGCA 1500
QY 1501 CATTTACCAATCACTGAAATTTGATGCGGATGTCAGCTTAAACGTAAGCATTTCC 1560
DB 1501 CATTTACCAATCACTGAAATTTGATGCGGATGTCAGCTTAAACGTAAGCATTTCC 1560
QY 1561 AGGCTCTGATTCATATCAATCCAAAGATTAAACATGCAAGATTTGGAAA 1620
DB 1561 AGGCTCTGATTCATATCAATCCAAAGATTAAACATGCAAGATTTGGAAA 1620
QY 1621 CACTTGGATGATTTGACGTACCGGATATGCTTCAAGATCTTCAATCAACCGT 1680
DB 1621 CACTTGGATGATTTGACGTACCGGATATGCTTCAAGATCTTCAATCAACCGT 1680
QY 1681 TTGAGTCAATCCCGTATTTAGCGCGGTATGAGGTGTTTATTAATAAAAGCTGC 1740
DB 1681 TTGAGTCAATCCCGTATTTAGCGCGGTATGAGGTGTTTATTAATAAAAGCTGC 1740
QY 1741 CGAAGTCTGGGAGCTTTTATTAACAGATATGCTTAAAGCGTTTGGGCAAG 1800
DB 1741 CGAAGTCTGGGAGCTTTTATTAACAGATATGCTTAAAGCGTTTGGGCAAG 1800
QY 1801 AATGCTCGGAGATGAGCAAAACACCAATTAAGTACTACCAACCACTTTGCT 1860
DB 1801 AATGCTCGGAGATGAGCAAAACACCAATTAAGTACTACCAACCACTTTGCT 1860
QY 1861 CTTACAGCCCAATGATGAGTACGACCTTTAATAGGCACTTGGCTAAGAAAG 1920
DB 1861 CTTACAGCCCAATGATGAGTACGACCTTTAATAGGCACTTGGCTAAGAAAG 1920
QY 1921 AATACGTTAATCAAGACGTAAGCATCAAGTAAAGCTTAAAGTGAACGAGCAATTTG 1980
DB 1921 AATACGTTAATCAAGACGTAAGCATCAAGTAAAGCTTAAAGTGAACGAGCAATTTG 1980
QY 1981 CAGAGCAACAGGCAATTCACAGAGACAGGCTTGGCGAGACAGACAGTAAATGA 2040
```



```

Db      1981 CAGAGCAACAGCGCAAACTCACACAGAGACAGCGGTTGGCGGAGAGACAGTAATACA 2040
QY      2041 AGTCCCAATTTTGGACCTTAAGTAAATGGTATGATTTTCAACGCACTTTCAACGCGCAGA 2100
Db      2041 AGTCCCAATTTTGGACCTTAAGTAAATGGTATGATTTTCAACGCACTTTCAACGCGCAGA 2100
QY      2101 GCCCAGAGAGAGAACCAATTAGGCTGGTGGTGGTGCATGAAAGATTGAATTAATACGCTAAC 2160
Db      2101 GCCCAGAGAGAGAACCAATTAGGCTGGTGGTGGTGCATGAAAGATTGAATTAATACGCTAAC 2160
QY      2161 CACTGTACTGTAAGCAATAACCGTGTAGTGGGCTCGGCAATTCGATTTTGTAGAAATTC 2220
Db      2161 CACTGTACTGTAAGCAATAACCGTGTAGTGGGCTCGGCAATTCGATTTTGTAGAAATTC 2220
QY      2221 AGGTGGCGGCGCAACCATCAACTCTTCATCAGTTTGGCCATCACCGTAAATGGCGACGA 2280
Db      2221 AGGTGGCGGCGCAACCATCAACTCTTCATCAGTTTGGCCATCACCGTAAATGGCGACGA 2280
QY      2281 AATGTCGCAATACCCCAATACGATTAAGTGCACACCCAGCAAGTATTACCAATACCGA 2340
Db      2281 AATGTCGCAATACCCCAATACGATTAAGTGCACACCCAGCAAGTATTACCAATACCGA 2340
QY      2341 AAGCTGGGTTTCAAGCCCTGTGTATGACAGGTTTGGTAATCGGTTTGATTAATCAAACC 2400
Db      2341 AAGCTGGGTTTCAAGCCCTGTGTATGACAGGTTTGGTAATCGGTTTGATTAATCAAACC 2400
QY      2401 TTTTATGCTCATATCGCAGTCGCAAGAGCGGCGAAGCGAGCGAGTGTGAATCTTTCTC 2460
Db      2401 TTTTATGCTCATATCGCAGTCGCAAGAGCGGCGAAGCGAGCGAGTGTGAATCTTTCTC 2460
QY      2461 TAAATATCCAAATCATATTTCTAGAGGTAGAAAGATGACACCGCGAGAAATTTAAAAA 2520
Db      2461 TAAATATCCAAATCATATTTCTAGAGGTAGAAAGATGACACCGCGAGAAATTTAAAAA 2520
QY      2521 ATCGTGGATGGTGGGCACTGGCGAAGAGCGGGAATCTTTTACGCGGATATGGCC 2580
Db      2521 ATCGTGGATGGTGGGCACTGGCGAAGAGCGGGAATCTTTTACGCGGATATGGCC 2580
QY      2581 AAGGCTGACGAGAGATTTGGTCAACAGTACCAATATTGGACCCATCACCATAGGAAT 2640
Db      2581 AAGGCTGACGAGAGATTTGGTCAACAGTACCAATATTGGACCCATCACCATAGGAAT 2640
QY      2641 CGCGGTTTAAACCGGTAACCCACCGGCAACGACCAATTAATGAAGTACCGTGCT 2700
Db      2641 CGCGGTTTAAACCGGTAACCCACCGGCAACGACCAATTAATGAAGTACCGTGCT 2700
QY      2701 TGAAGATTGAATCAGTGGCGTGGCCACTAAACCAATCATCATCTGCAATTTGGGTGGA 2760
Db      2701 TGAAGATTGAATCAGTGGCGTGGCCACTAAACCAATCATCATCTGCAATTTGGGTGGA 2760
QY      2761 AGCAATTTCAATTAAGACTTTGGCTTGAATGCGCGGTTGGCCATTTAAACCGCTCCGAC 2820
Db      2761 AGCAATTTCAATTAAGACTTTGGCTTGAATGCGCGGTTGGCCATTTAAACCGCTCCGAC 2820
QY      2821 CATCGGATCGCAAGAGATTAATTAACGATGAAAGCCAAAGTTTGCCCAAGTAGGCC 2880
Db      2821 CATCGGATCGCAAGAGATTAATTAACGATGAAAGCCAAAGTTTGCCCAAGTAGGCC 2880
QY      2881 TTTTGGTGTGACGCAAAATCGCGCTGACG 2909
Db      2881 TTTTGGTGTGACGCAAAATCGCGCTGACG 2909

```

```

RESULT 2
AAZ40648
ID      AAZ40648 standard; DNA; 1222 BP.
XX
XX      AAZ40648;
AC      08-MAR-2000 (first entry)
XX
XX      Nucleotide sequence of V. cholerae thya gene 3' flanking region.
DE
XX

```

```

KM      Vibrio cholerae; thya-negative strain; Deltathya; thya gene; vaccine;
KM      cholera; antimicrobial therapy; ss.
XX
XX      Vibrio cholerae.
OS
XX      W09961634-A1.
PN
XX      02-DEC-1999.
PD
XX      21-MAY-1999; 99WO-Ep003509.
PF
XX      26-MAY-1998; 98SE-00001852.
PR
XX      (SBLV-) SBL VACCIN AB.
PA
XX      Carlin N, Lebens MR;
PI
XX      WPI; 2000-062719/05.
DR
XX
XX      New Vibrio cholerae strain defective in the thya gene, for use in
XX      vaccines and for recombinant protein production.
PT
XX      Claim 11; Fig 3; 42pp; English.
PS
XX
XX      The invention provides a method for producing a Vibrio cholerae thya-
XX      negative strain which is Deltathya strain lacking thya gene functions.
XX      The method comprises site-directed mutagenesis of the V. cholerae
XX      chromosome to delete and/or insert nucleotides at the thya locus; The V.
XX      cholerae thya-negative strains are used: for overproduction of
XX      recombinant proteins; and in vaccines to prevent or treat cholera (or
XX      other diseases if engineered to express the appropriate proteins). The
XX      thya gene is also useful for insertion of foreign genes, in a selective
XX      and site-specific manner, and the proteins expressed by the thya gene or
XX      by its 5'-flanking region, are useful in research and as targets for
XX      antimicrobial therapy. When used for recombinant protein production, V.
XX      cholerae provides high yields with secretion of products into the culture
XX      medium for ease of subsequent recovery. The thya-negative strain can be
XX      maintained by thymine complementation, eliminating the need for
XX      antibiotic selection. The present sequence represents the nucleotide
XX      sequence of the V. cholerae thya gene 3' flanking region
XX
SQ      Sequence 1222 BP; 373 A; 293 C; 289 G; 267 T; 0 U; 0 Other;
Query Match      42.0%; Score 1222; DB 3; Length 1222;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1688 TAAATCCGATTCAGGCGGATATGCTTGAATGGGTTTATATAAAAAGCTCCGAAAGT 1747
Db      1 TAAATCCGATTCAGGCGGATATGCTTGAATGGGTTTATATAAAAAGCTCCGAAAGT 60
QY      1748 CGGAGCTTTTATATACAGATGATGCTTAACGCTTAACGGTTAGGGCAAGATGCTG 1807
Db      61 CGGAGCTTTTATATACAGATGATGCTTAACGCTTAACGGTTAGGGCAAGATGCTG 120
QY      1808 CGGAGATGACGACAAACACCAATTAAGTCAACACCAACATTTTGTCTTACAA 1867
Db      121 CGGAGATGACGACAAACACCAATTAAGTCAACACCAACATTTTGTCTTACAA 180
QY      1868 GCCCAATTGATGATGAGTACGACACCTTAATAGGCGGTTGGGTAAGAAAGAAATACCG 1927
Db      181 GCCCAATTGATGATGAGTACGACACCTTAATAGGCGGTTGGGTAAGAAAGAAATACCG 240
QY      1928 TAAATCAAGACCGTAGCATCAAGTTAAAGCTTAAGTGCACGACCGCAATTTGCAGAGCA 1987
Db      241 TAAATCAAGACCGTAGCATCAAGTTAAAGCTTAAGTGCACGACCGCAATTTGCAGAGCA 300
QY      1988 AACACGGCAAAATCAACAGAGAGACGCGTTGGCGGACGAGACGATTAATACAGTGGCA 2047
Db      301 AACACGGCAAAATCAACAGAGAGACGCGTTGGCGGACGAGACGATTAATACAGTGGCA 360
QY      2048 ATGTTCGACACTAAGTAATGGTATGATTTCAACGCACTTTCAACGACGCGCAGGCCACG 2107

```

D	b		361	ATGTTGGCACTAAGTAAGTAATGGGTAAATTTCACGCACCTTTCACGACGCCAGAGCCAGC	420
Q	y		2108	AGAGGAACCAATTAGGCTGGTGTGGTGCATGAAGATTGAACTAATTACCGTAACCACTGTA	2167
D	b		421	AGAGGAACCAATTAGGCTGGTGTGGTGCATGAAGATTGAACTAATTACCGTAACCACTGTA	480
Q	y		2168	CCGTAAGCAATPACCGGTGATGAGGGGCTCGGCGCAATCGCAATTTTGTAGAATTTCAGTGC	2227
D	b		481	CTTGAAGCAATPACCGGTGATGAGGGGCTCGGCGCAATCGCAATTTTGTAGAATTTCAGTGC	540
Q	y		2228	CGGCGCAACCATCAAACTCTTTCATCAGTTTGGCCATCACCGTAATGGGACGAAAAATGTC	2287
D	b		541	CGGCGCAACCATCAAACTCTTTCATCAGTTTGGCCATCACCGTAATGGGACGAAAAATGTC	600
Q	y		2288	GCATTAACCATATGATTAAGTGGCAGCACCAACGAAGATTATACCAATACGGAAGCTGG	2347
D	b		601	GCAATTAACCATATGATTAAGTGGCAGCACCAACGAAGATTATACCAATACGGAAGCTGG	660
Q	y		2348	GTTTCAAGCCCTGTGATGACAGGTTTGGTAATCGGTTTGAATAAATCAAACTTTTCATG	2407
D	b		661	GTTTCAAGCCCTGTGATGACAGGTTTGGTAATCGGTTTGAATAAATCAAACTTTTCATG	720
Q	y		2408	CTCATATCGCAGTCCGAGACGACAGGAGGGGAAACGAGGCAAGTGTGAGACTTTCTTAAATG	2467
D	b		721	CTCATATCGCAGTCCGAGACGAGGAGGGGAAACGAGGCAAGTGTGAGACTTTCTTAAATG	780
Q	y		2468	CCAACACATCATTTTGAAGGTAGGAGATCAGCACCGGAGAGAATTTGAAAAATCGTGG	2527
D	b		781	CCAACACATCATTTTGAAGGTAGGAGATCAGCACCGGAGAGAATTTGAAAAATCGTGG	840
Q	y		2528	ATGTTGGCACTGGCGGAAAGCACCGCGGAAACTCTTCTTTACAGCGCATATGGCCAAAGCTG	2587
D	b		841	ATGTTGGCACTGGCGGAAAGCACCGCGGAAACTCTTCTTTACAGCGCATATGGCCAAAGCTG	900
Q	y		2588	ACGAGAGATTGTGTACACAGTAGTACCAATATTGGCACCACATAGAAATCGCGTT	2647
D	b		901	ACGAGAGATTGTGTACACAGTAGTACCAATATTGGCACCACATAGAAATCGCGTT	960
Q	y		2648	TCAAACCGGTAACCCACCGGCAACGAGACCAACATATAGAAGTACCGTGTGAGGAT	2707
D	b		961	TCAAACCGGTAACCCACCGGCAACGAGACCAACATATAGAAGTACCGTGTGAGGAT	1020
Q	y		2708	TGAATCAATGCGCGTTGCCACTTAACCAATCATCATCTGTGAATTGCGTGGGAAAGCAAT	2767
D	b		1021	TGAATCAATGCGCGTTGCCACTTAACCAATCATCATCTGTGAATTGCGTGGGAAAGCAAT	1080
Q	y		2768	TCAATATAGACTTTGGCTTGAATCGCGGGTTGCCATTTTAAACCGCTGCCACCATCGCG	2827
D	b		1081	TCAATATAGACTTTGGCTTGAATCGCGGGTTGCCATTTTAAACCGCTGCCACCATCGCG	1140
Q	y		2828	ACTGCAAGAAAGTAGTAATTAACAGCATGAAGCAAGTTTGGCCAAAGTAAAGGCTTTCGTG	2887
D	b		1141	ACTGCAAGAAAGTAGTAATTAACAGCATGAAGCAAGTTTGGCCAAAGTAAAGGCTTTCGTG	1200
Q	y		2888	GTCAGCGCAATCGCGCTGCGAG	2909
D	b		1201	GTCAGCGCAATCGCGCTGCGAG	1222
RESULT 3					
ACA53067/c					
ID ACA53067 standard; DNA; 1149 BP.					
XX	AC	ACA53067;			
XX	DT	19-JUN-2003 (first entry)			
XX	DE	Prokaryotic essential gene #34724.			
XX	KM	Antisense; ds; prokaryotic essential gene; cell proliferation;			
XX	XX	drug design; gene.			
OS	Vibrio cholerae.				

QY 1903 CAGTTCCGCTAAGAAAGAAATACCGTAATCAAGACCGTAGCCATCAAGTTAAAGCTTAA 1962
DB 1029 CAGTTCCGCTAAGAAAGAAATACCGTAATCAAGACCGTAGCCATCAAGTTAAAGCTTAA 970
QY 1963 GTGCACCAAGCGCAATTTGCAAGCAAAACCGGCAAACTCAGAGAGACCGGTTGCGGC 2022
DB 969 GTGCACCAAGCGCAATTTGCAAGCAAAACCGGCAAACTCAGAGAGACCGGTTGCGGC 910
QY 2023 GAGCAGAGAGTAATCAAGTCCCAATGTTCCGACCTAAGTAATGGTGTGATTTCAG 2082
DB 909 GAGCAGAGAGTAATCAAGTCCCAATGTTCCGACCTAAGTAATGGTGTGATTTCAG 850
QY 2083 CACTTTCAGCAGCAGCAGACCGCAGAGAGAAACATTAAGCTGTTGTGTGATGAAGA 2142
DB 849 CACTTTCAGCAGCAGCAGACCGCAGAGAGAAACATTAAGCTGTTGTGTGATGAAGA 790
QY 2143 TTGAACTAATACCGTAACCTGTATCTGAAGCAATACCGTGTATGCGGCTCGCCAT 2202
DB 789 TTGAACTAATACCGTAACCGTAACCTGTATCTGAAGCAATACCGTGTATGCGGCTCGCCAT 730
QY 2203 CCGATTTTGTAGAAATTTCACTGCGCGGCAACCACTCAAACTCTTCATCAATTTGCCAT 2262
DB 729 CCGATTTTGTAGAAATTTCACTGCGCGGCAACCACTCAAACTCTTCATCAATTTGCCAT 670
QY 2263 CACCGTAATGCGCAGCAAAATGCTGCGCAATACCCAAATAGATTAAGTGCACACCGCA 2322
DB 669 CACCGTAATGCGCAGCAAAATGCTGCGCAATACCCAAATAGATTAAGTGCACACCGCA 610
QY 2322 AGTATTAACCAATACCGAAAGTGGGTTTCAAGCCCTGAGAGAGAGTTGGTATAGG 2382
DB 609 AGTATTAACCAATACCGAAAGTGGGTTTCAAGCCCTGATTAAGAGAGTTGGTATAGG 550
QY 2383 TTTGATTAATCAAAACCTTTCATGCTCATATGCGCAATGCGCAAGCGGCAAGCG 2442
DB 549 TTTGATTAATCAAAACCTTTCATGCTCATATGCGCAATGCGCAAGCGGCAAGCG 490
QY 2443 CCAAGTGTAGACTTTCTTAAATGCGCAACATCATTTCTAAGAGTAGAAGATAGAC 2502
DB 489 CCAAGTGTAGACTTTCTTAAATGCGCAACATCATTTCTAAGAGTAGAAGATAGAC 430
QY 2503 CCGGAGAAATGAGAAATGCGGATGCTGCGCACTGCGGAAAGCAGCGCAAACTTTC 2562
DB 429 CCGGAGAAATGAGAAATGCGGATGCTGCGCACTGCGGAAAGCAGCGCAAACTTTC 370
QY 2563 TTTAAGCGCATATGCGCAAGCTGACAGAGATTTGTCACAGTAGTACCAATTTGGC 2622
DB 369 TTTAAGCGCATATGCGCAAGCTGACAGAGATTTGTCACAGTAGTACCAATTTGGC 310
QY 2623 ACCCATCAACATAGGAATGCGGTTTCAACCGGTAAACCGGCAAGCAGACCAAT 2682
DB 309 ACCCATCAACATAGGAATGCGGTTTCAACCGGTAAACCGGCAAGCAGACCAAT 250
QY 2683 AATAGAGTACCGTCTTGAAGATTTGAATCAAGTCCGCTTGCACATAACCAATCA 2742
DB 249 AATAGAGTACCGTCTTGAAGATTTGAATCAAGTCCGCTTGCACATAACCAATCA 190
QY 2743 TCTGTCAATTTGGGTGGGAAACCAATTTCAATTAAGACTTTGGCTGTATGCGGCTGCCA 2802
DB 189 TCTGTCAATTTGGGTGGGAAACCAATTTCAATTAAGACTTTGGCTGTATGCGGCTGCCA 130
QY 2803 TTTAAACCGCTGCGCAGCATGCGCACTGCAAGAGTAGTAATACAGATGAAGCCAA 2862
DB 129 TTTAAACCGCTGCGCAGCATGCGCACTGCAAGAGTAGTAATACAGATGAAGCCAA 70
QY 2863 GTTTGCCCAAGTAGGCTTTGTGTGTCAGGAAATCGGCGCTGCAG 2909
DB 69 GTTTGCCCAAGTAGGCTTTGTGTGTCAGGAAATCGGCGCTGCAG 23

RESULT 4
AAZ24801
ID AAZ24801 standard; DNA; 852 BP.
XX

AC AAZ24801;
XX
DT 07-DEC-1999 (first entry)
XX
DE Vibrio cholerae thA coding region sequence.
XX
KW Wild type; thA; enzyme; thymidilate synthase; non-toxicogenic; mutation;
XX attenuation; immunization; cholera; hemagglutinin protease; biosafety;
XX auxotrophism; replication; vaccine; ss.
XX
OS Vibrio cholerae.
XX
PN M0935271-A2.
XX
PD 15-JUL-1999.
XX
PE 30-DEC-1998; 98MO-CU000008.
XX
PR 30-DEC-1997; 97CU-00000142.
XX
PA (NAIN-) CENT NACIONAL INVESTIGACIONES CIENTIFICA.
XX
PI Campos Gomez J, Fando Galzada RA, Rodriguez Gonzalez BL;
XX Leon Perez TY, Valle Diaz E, Silva Cabrera AJ, Benitez Robles JA;
XX WPI; 1999-430398/36.
XX
DR P-PSDB; AAY26895.
XX
PT Producing strains of Vibrio cholerae with inactivated gene for
XX hemagglutinin protease, useful in vaccines against cholera.
XX
PS Claim 18; Page 30; 30pp; Spanish.
XX
SS This sequence represents the wild type coding region of the thA gene
XX from Vibrio cholerae, which encodes the enzyme thymidilate synthase. The
XX invention relates to the production, from a non-toxicogenic strain of V.
XX cholerae, of attenuated strains of V. cholerae suitable for immunization
XX against cholera. The attenuation comprises inactivating the gene for
XX hemagglutinin protease (HP), either by deletion, insertion or some other
XX defined and irreversible genetic manipulation. Additional biosafety of
XX the attenuated strains is generated by mutating the thA gene, resulting
XX in auxotrophic mutants unable to replicate in the environment. The new
XX strains are used to produce anticholera vaccines
XX
SQ Sequence 852 BP; 217 A; 177 C; 216 G; 242 T; 0 U; 0 Other;
Query Match 29.0%; Score 842.4; DB 2; Length 852;
Best Local Similarity 99.3%; Pred. No. 1.7e-257;
Matches 846; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 839 GTGAAACAGATTTTAAATCTTTGTACGCGCATGTGATTAAGGTGTTGGTTGAAAT 898
DB 1 GTGAAACAGATTTTAAATCTTTGTACGCGCATGTGATTAAGGTGTTGGTTGAAAT 60
QY 899 GAAAGAACGCGCAAGCGTTGTTGACTGTGATTAATGCCATTGACATGTGCGC 958
DB 61 GAAAGAACGCGCAAGCGTTGTTGACTGTGATTAATGCCATTGACATGTGCGC 120
QY 959 AACATAGTCTTCTCTTGAAGTACACGCAAGAGTTTGGAAAGTGCCTAGCCGAG 1018
DB 121 AACATAGTCTTCTCTTGAAGTACACGCAAGAGTTTGGAAAGTGCCTAGCCGAG 180
QY 1019 TTGCTGGCTAATTTGTTGTTAGTAAGTAAGCGGCGGATTTTGGCAATTAAGTACCAA 1078
DB 181 TTGCTGGCTAATTTGTTGTTAGTAAGTAAGCGGCGGATTTTGGCAATTAAGTACCAA 240
QY 1079 ACCTGGATGCTAATGCGCAATTTAAACCAAGCATGCTCAACATCTCTTAACGTTAAAGT 1138
DB 241 ACCTGGATGCTAATGCGCAATTTAAACCAAGCATGCTCAACATCTCTTAACGTTAAAGT 300
QY 1139 GAGGATGACATGGGACCGGTGATGTTGAGGTGAGGTTAGGCTTAAAGCTGATAGT 1198
DB 301 GAGGATGACATGGGACCGGTGATGTTGAGGTGAGGTTAGGCTTAAAGCTGATAGT 360

QY 1199 GGTCAATTTGACCAAGTTGAAAAAGATTGTTGATTTGAGCCGTGGCGTTGATGACCGA 1258
DB 361 GGTCAATTTGACCAAGTTGAAAAAGATTGTTGATTTGAGCCGTGGCGTTGATGACCGA 420
QY 1259 GGTGAATTTCTTACTCTTCAACATCCGGGTGAAATTTCAATGGGGTGGTGGCCCTTGC 1318
DB 421 GGTGAATTTCTTACTCTTCAACATCCGGGTGAAATTTCAATGGGGTGGTGGCCCTTGC 480
QY 1319 ATGTACAGCCATCTTTCATTTGCTGCTGGGGATACCTTGTATCTCAACAGTACGAGCT 1378
DB 481 ATGTACAGCCATCTTTCATTTGCTGCTGGGGATACCTTGTATCTCAACAGTACGAGCT 540
QY 1379 TCATGTGATGTCCTTGGGGTGGTGAATTTCAACAGTGGTGGTGAATTTGCTGCTGG 1438
DB 541 TCATGTGATGTCCTTGGGGTGGTGAATTTCAACAGTGGTGGTGAATTTGCTGCTGG 600
QY 1439 CTGATGTCACAGATTCACAGGAAAAAGCCGGCTTGGCTATCAACAAGTGTCAATGCG 1498
DB 601 CTGATGTCACAGATTCACAGGAAAAAGCCGGCTTGGCTATCAACAAGTGTCAATGCG 660
QY 1499 CACATTTACCAAGATTAATCTCGATTTGATGCGGATGTCAGCTTAAACGTGACCATTC 1558
DB 661 CACATTTACCAAGATTAATCTCGATTTGATGCGGATGTCAGCTTAAACGTGACCATTC 720
QY 1559 CCAGCCGCTCAGTTCCATATTCATCCAAAGATTAAACACTGACGATTTGGAACCTTGG 1618
DB 721 CCAGCCGCTCAGTTCCATATTCATCCAAAGATTAAACACTGACGATTTGGAACCTTGG 780
QY 1619 GTCACTTTGGATGATTTTGAACGTACCGGATTAATCAAGTTCATTTCAATACCGG 1678
DB 781 GTCACTTTGGATGATTTTGAACGTACCGGATTAATCAAGTTCATTTCAATACCGG 840
QY 1679 TTTTCAGTCTAA 1690
DB 841 TTTTCAGTCTAA 852

RESULT 5
ACAS3066
ID ACAS3066 standard; DNA; 852 BP.
AC 19-JUN-2003 (first entry)
DT 19-JUN-2003 (first entry)
DE Prokaryotic essential gene #34723.
XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;
KM drug design; gene.
XX
OS Vibrio cholerae.
XX
PN W020027783-A2.
PD 03-OCT-2002.
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 26-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0142923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zvekind JW;
PI Wall D, Treweek JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
DR P-PSDB; ABU49196.
XX

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids, required for cellular proliferation or
PT isolate candidate molecules for rational drug discovery programs.
PS Claim 14; SEQ ID NO 40936; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 613 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcr_sequences
XX
SQ Sequence 852 BP; 217 A; 177 C; 216 G; 242 T; 0 U; 0 Other;
Query Match 29.0%; Score 842.4; DB 7; Length 852;
Best Local Similarity 99.3%; Pred. No. 1.7e-257;
Matches 846; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 839 GTGAACAGTATTAGATCTTTGTCAAGCCGATGTCGATCAAGGTTGGGTGAAT 898
DB 1 GTGAACAGTATTAGATCTTTGTCAAGCCGATGTCGATCAAGGTTGGGTGAAT 60
QY 899 GAACGAACGGGCAAGGTTGTGACGTGATTAATGCGGATTTGGAAGCTGCGTGAAG 958
DB 61 GAACGAACGGGCAAGGTTGTGACGTGATTAATGCGGATTTGGAAGCTGCGTGAAG 120
QY 959 AACATCAGTTTCTTCTAGTACTACAACGAGATTTTGAAGCTGCGTGAAG 1018
DB 121 AACATCAGTTTCTTCTAGTACTACAACGAGATTTTGAAGCTGCGTGAAG 180
QY 1019 TTGCTGGCTATATTTGCTGTTAGATTAATGCGGATTTTGAAGCTGCGTGAAG 1078
DB 181 TTGCTGGCTATATTTGCTGTTAGATTAATGCGGATTTTGAAGCTGCGTGAAG 240
QY 1079 AACTGGATCTATATGCAATTTAAACCAAGCATGCTCAACATCTTAACCTTAAGGT 1138
DB 241 AACTGGATCTATATGCAATTTAAACCAAGCATGCTCAACATCTTAACCTTAAGGT 300
QY 1139 GAGGATACATGGGACCGCTGTATGTTTCAAGGTTAGAGCTTGGGCTTAAGCTGATGT 1198
DB 301 GAGGATACATGGGACCGCTGTATGTTTCAAGGTTAGAGCTTGGGCTTAAGCTGATGT 360
QY 1199 GGTCAATTTGACCAAGTTGAAAAAGATTGTTGATTTGAGCCGTGGCGTTGATGACCGA 1258
DB 361 GGTCAATTTGACCAAGTTGAAAAAGATTGTTGATTTGAGCCGTGGCGTTGATGACCGA 420
QY 1259 GGTGAATTTCTTACTCTTCAACATCCGGGTGAAATTTCAATGGGGTGGTGGCCCTTGC 1318

DB 421 GGTGAATTTCTTAACCTTCAATCCGGGTGATTTTCATGGGGTGTGGCCCTTGC 480
QY 1319 ATGTAAGCCATCATTTTTCATTGCTGGGGATACCTTGTATCTCAAGTAAGCGGT 1378
DB 481 ATGTAAGCCATCATTTTTCATTGCTGGGGATACCTTGTATCTCAAGTAAGCGGT 540
QY 1379 TCATGTGATGTGCTGGGGTGGTGAATTTCAACATGGTGCAGCTTATGTCTTCCG 1438
DB 541 TCATGTGATGTGCTGGGGTGGTGAATTTCAACATGGTGCAGCTTATGTCTTCCG 600
QY 1439 CTGATGCAAGATCAAGGCAAAAGCGGGCTTGCGTATCAACAGATGTCATGCG 1498
DB 601 CTGATGCAAGATCAAGGCAAAAGCGGGCTTGCGTATCAACAGATGTCATGCG 660
QY 1499 CACATTTACCAAGATCAAGTCAATTTGATGCGCGATGTGCACTTAAACGTAGCATTC 1538
DB 661 CACATTTACCAAGATCAAGTCAATTTGATGCGCGATGTGCACTTAAACGTAGCATTC 720
QY 1559 CAGAGGCGCTCAGTTCATATCATTCGAAAGATTAAACATGCAAGATTTGGAACCTGG 1618
DB 721 CAGAGGCGCTCAGTTCATATCATTCGAAAGATTAAACATGCAAGATTTGGAACCTGG 780
QY 1619 GTCACTTTGATGATTTTTCAGTCAACCGGATTCAGTTCCAGATCTTATTCATACCG 1678
DB 781 GTCACTTTGATGATTTTTCAGTCAACCGGATTCAGTTCCAGATCTTATTCATACCG 840
QY 1679 TTTTCAGTCTAA 1690
DB 841 TTTTCAGTCTAA 852

RESULT 6
AAZ40647
ID AAZ40647 standard; DNA; 838 BP.
XX
XX AAZ40647;
XX
XX 08-MAR-2000 (first entry)
XX
DE Nucleotide sequence of V. cholerae thya gene 5' flanking region.
XX
XX Vibrio cholerae; thya-negative strain; Deltathya; thya gene; vaccine;
XX
XX cholera; antimicrobial therapy; ss.
XX
XX Vibrio cholerae.
XX
XX MO961634-A1.
XX
XX 02-DEC-1999.
XX
XX 21-MAY-1999; 99WO-EP003509.
XX
XX 26-MAY-1998; 98SE-00001852.
XX
XX (SBLV-) SBL VACCIN AB.
XX
XX Carlin N, Lebens MR;
XX
XX WPI; 2000-062719/05.
XX
XX P-PsDB; AAY59127.
XX
XX New Vibrio cholerae strain defective in the thya gene, for use in
XX
XX vaccines and for recombinant protein production.
XX
XX Claim 10; Fig 2; 42pp; English.
XX
XX The invention provides a method for producing a Vibrio cholerae thya-
XX
XX negative strain which is Deltathya strain lacking thya gene functions.
XX
XX The method comprises site-directed mutagenesis of the V. cholerae
XX
XX chromosome to delete and/or insert nucleotides at the thya locus; the V.
XX
XX cholerae thya-negative strains are used; for overproduction of
XX
XX recombinant proteins; and in vaccines to prevent or treat cholera (or

CC other diseases if engineered to express the appropriate proteins). The
CC thya gene is also useful for insertion of foreign genes, in a selective
CC and site-specific manner, and the proteins expressed by the thya gene or
CC by its 5'-flanking region, are useful in research and as targets for
CC antimicrobial therapy. When used for recombinant protein production, V.
CC cholerae provides high yields with secretion of products into the culture
CC medium for ease of subsequent recovery. The thya-negative strain can be
CC maintained by thymine complementation, eliminating the need for
CC antibiotic selection. The present sequence represents the nucleotide
CC sequence of the V. cholerae thya gene 5' flanking region
XX
XX Sequence 838 BP; 145 A; 176 C; 232 G; 285 T; 0 U; 0 Other;
SQ
Query Match 28.8%; Score 838; DB 3; Length 838;
Best Local Similarity 100.0%; Pred. No. 4.4e-256;
Matches 838; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGAAAGTTTGTATAGCTTCAGGATATCTGCAAGTTTCCCAATATGACCCCGATTTGTT 60
DB 1 GAGAAAGTTTGTATAGCTTCAGGATATCTGCAAGTTTCCCAATATGACCCCGATTTGTT 60
QY 61 TTGATGCGGCGCTCAGGATGCGGCTGATGCTGATGATTTGGTGGTTCCTTT 120
DB 61 TTGATGCGGCGCTCAGGATGCGGCTGATGCTGATGATTTGGTGGTTCCTTT 120
QY 61 TTGATGCGGCGCTCAGGATGCGGCTGATGCTGATGATTTGGTGGTTCCTTT 120
DB 61 TTGATGCGGCGCTCAGGATGCGGCTGATGCTGATGATTTGGTGGTTCCTTT 120
QY 121 TGCTATGTGTTGCGCAATGCGGAGCGGATGCGGCGGAGTGTGACGCGGTAGCA 180
DB 121 TGCTATGTGTTGCGCAATGCGGAGCGGATGCGGCGGAGTGTGACGCGGTAGCA 180
QY 121 TGCTATGTGTTGCGCAATGCGGAGCGGATGCGGCGGAGTGTGACGCGGTAGCA 180
DB 121 TGCTATGTGTTGCGCAATGCGGAGCGGATGCGGCGGAGTGTGACGCGGTAGCA 180
QY 181 AGTCTGCACTTGTATTCGCGCGCTTTTATGATGTATGATGATGCGGCGGATGTTA 240
DB 181 AGTCTGCACTTGTATTCGCGCGCTTTTATGATGTATGATGATGCGGCGGATGTTA 240
QY 181 AGTCTGCACTTGTATTCGCGCGCTTTTATGATGTATGATGATGCGGCGGATGTTA 240
DB 181 AGTCTGCACTTGTATTCGCGCGCTTTTATGATGTATGATGATGCGGCGGATGTTA 240
QY 241 TGTATCTTCAATTTTGTATGCTGCTGCTGCAACCTTATTTATTTCAAGTGTG 300
DB 241 TGTATCTTCAATTTTGTATGCTGCTGCTGCAACCTTATTTATTTCAAGTGTG 300
QY 301 GACTGCGGCGATGCTTTCACAGCGGCTTATGCGGTGATCAACCGCATGTTCTGTA 360
DB 301 GACTGCGGCGATGCTTTCACAGCGGCTTATGCGGTGATCAACCGCATGTTCTGTA 360
QY 301 GACTGCGGCGATGCTTTCACAGCGGCTTATGCGGTGATCAACCGCATGTTCTGTA 360
DB 301 GACTGCGGCGATGCTTTCACAGCGGCTTATGCGGTGATCAACCGCATGTTCTGTA 360
QY 361 TGCGCGTAAACCAACAGCAACCTTCTGCGTGGCGGATTTGTTGCCCTTAGTGCC 420
DB 361 TGCGCGTAAACCAACAGCAACCTTCTGCGTGGCGGATTTGTTGCCCTTAGTGCC 420
QY 361 TGCGCGTAAACCAACAGCAACCTTCTGCGTGGCGGATTTGTTGCCCTTAGTGCC 420
DB 361 TGCGCGTAAACCAACAGCAACCTTCTGCGTGGCGGATTTGTTGCCCTTAGTGCC 420
QY 421 ATTGCGTTTGGGATGAGAGGTATCGGTAACTTTATGATATGAACTTTGGGACGAGT 480
DB 421 ATTGCGTTTGGGATGAGAGGTATCGGTAACTTTATGATATGAACTTTGGGACGAGT 480
QY 481 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTCGCGCATGCGGCGCATCTTC 540
DB 481 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTCGCGCATGCGGCGCATCTTC 540
QY 481 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTCGCGCATGCGGCGCATCTTC 540
DB 481 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTCGCGCATGCGGCGCATCTTC 540
QY 541 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTCGCGCATGCGGCGCATCTTC 541
DB 541 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTCGCGCATGCGGCGCATCTTC 541
QY 601 TGTAAACCTTGTGCTGCTAGGACGCTATCGGACTGTTTATGCTGATACGTAAT 660
DB 601 TGTAAACCTTGTGCTGCTAGGACGCTATCGGACTGTTTATGCTGATACGTAAT 660
QY 601 TGTAAACCTTGTGCTGCTAGGACGCTATCGGACTGTTTATGCTGATACGTAAT 660
DB 601 TGTAAACCTTGTGCTGCTAGGACGCTATCGGACTGTTTATGCTGATACGTAAT 660
QY 661 CCGCTTCTTGTGAATACGCTGAGCGGATGAGCGGATGAGCGGATGAGCGGATGAG 720
DB 661 CCGCTTCTTGTGAATACGCTGAGCGGATGAGCGGATGAGCGGATGAGCGGATGAG 720
QY 721 CATTCAATGCGGCAAAATCTCTCTTACCTATGCTATCATCGGTAATTTGATGATGT 780
DB 721 CATTCAATGCGGCAAAATCTCTCTTACCTATGCTATCATCGGTAATTTGATGATGT 780
QY 781 TTGCTTACAGAGCGGCTTTGTATCAAGACCGGTGACGCAAAATAGGATAGT 838
DB 781 TTGCTTACAGAGCGGCTTTGTATCAAGACCGGTGACGCAAAATAGGATAGT 838

```
RESULT 7
Continuation (10 of 19) of AAT42063 from base 900001 (Haemophilus influenzae complete ge
WP Sequence Split into 19 fragments LOCUS AAT42063 Accession AAT42063
WP Fragment Name Begin End
WP AAT42063_00 1 110000
WP AAT42063_01 100001 210000
WP AAT42063_02 200001 310000
WP AAT42063_03 300001 410000
WP AAT42063_04 400001 510000
WP AAT42063_05 500001 610000
WP AAT42063_06 600001 710000
WP AAT42063_07 700001 810000
WP AAT42063_08 800001 910000
WP AAT42063_09 900001 1010000
WP AAT42063_10 1000001 1110000
WP AAT42063_11 1100001 1210000
WP AAT42063_12 1200001 1310000
WP AAT42063_13 1300001 1410000
WP AAT42063_14 1400001 1510000
WP AAT42063_15 1500001 1610000
WP AAT42063_16 1600001 1710000
WP AAT42063_17 1700001 1810000
WP AAT42063_18 1800001 1830121

Query Match 25 6%; Score 744.2; DB 2; Length 110000;
Best Local Similarity 66.7%; Pred. No. 7.9e-225;
Matches 1134; Conservative 0; Mismatches 538; Indels 27; Gaps 4;

QY 2 AGAAGGTTGTTATGCTCAGGGTTATCTGCAAGTTCCCAATATGACCCCGTATTGTT 61
DB 58086 AATAGGAAATATGATGATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 58145
QY 62 TCGATGGC-----CCTAGCGCGCGCTGATAGGCTGATAGGCTGATAGGCTGAT 115
DB 58146 ACCCTGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 58205
QY 116 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 175
DB 58206 GTTTTTCAGCGTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 58265
QY 176 GAGCAAGCTCTGACTGCTTATTCGCGGCTTTTATGATGATGATGATGATGATGAT 235
DB 58266 GATCAAGTGTATGCTTACTTTCATGCAAGGTTTATGAGGCTTATGAGGCTTAT 58325
QY 236 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 295
DB 58326 GGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 58385
QY 296 GTGTGACTGCGCGATGCTTCACGCGGCTTATGATGATGATGATGATGATGATGAT 355
DB 58386 GTTGGGAAGTGAATGCTGCTCCACGGTGGCTTATGATGATGATGATGATGAT 58445
QY 356 TGGTATGCGGTTAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 415
DB 58446 TGAACATCTTATTTCTCAAAAAGTATTTTGGCAACGGCTGATTTGTTGGCTT 58505
QY 416 GTGCAATGCGTTTGGGATGGAAGCTATCGTAATGATGATGATGATGATGATGAT 475
DB 58506 ATTCGCTTGGTTAGTTAGTTAGGCAAGATTTGTAATTTCAATTTCAATTTCA 58565
QY 476 CGAGTAAGGATGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 532
DB 58566 CGGAAACGAATGCTTGGGCAATGATTTTCCGAATGATCTTCTTACTGCTGCT 58625
QY 533 CATCCTTCAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 592
DB 58626 CATCCTTCAGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 58685
QY 593 TGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 652
DB 58686 ATTTTATTAATAAACAAGCTCAATGAGCTTCTGTTGCAAGTTTATTTCTAAT 58745

QY 653 GGTACATTCGCTTCTTGTGGAATACGTCGAGCAGATGCTCAATGGCTGTTT 712
DB 58746 GGGCTTCTTCCGTTTATTTATGGAATATGTCGGAACCTGAAG-----TTGAAAATTT 58795
QY 713 GGTGCTTCAATTCATAGGGGCAAAATCCCTTACCTTACCTTATGATGATGATGATGAT 772
DB 58800 TTTGGATTTATACACAGGCAAGCCCTTTGCTTCCGATGATGATGATGATGATGAT 58859
QY 773 ATGATGCTTGTGTTTCAAGCGGCTTGTATCAAGCCGTATGACGCAAAATAGGCT 832
DB 58860 ATTAGCTTGGGCTTATTCAGCC-----AAAAGTCCGCTCAATAAATAGAG 58907
QY 833 AGTAGGTGAACAAGTATTTAGATCTTTGTCAAGGATGCTGATCAAGGCTGTTGGCT 892
DB 58908 ATTTTATGAGCAATATCTTGAAGCTTTGTGCGGCAATTTTGTGTAAGGGAATAGCT 58967
QY 893 GAAATGACCAACGCGCAAGCGCTTGTGATGATGATGATGATGATGATGATGATGAT 952
DB 58968 GCTATGACGTAAGGTAAGCATGCTCAGCGCTATTAATGCAATTTAGATGATGAT 59027
QY 953 GTGGGCAACATCAGTTTCTCTAGTACCTACACGCAAGGTTTGGAAAGCTGCGTA 1012
DB 59028 GTGGGCAATATCAATTTCCGCTGATTAATCCGTAAGGTAATGGAAGGCGGAT 59087
QY 1013 GCCAGTTGCTCGCTATATTTCTGTTAGCATATGCGCGGATTTTCCCAATTAGCT 1072
DB 59088 GCTGAATTTTATGATTAATTTGCTGATATGATGATGATGATGATGATGATGATGAT 59147
QY 1073 ACCAAATCTGGGATGCTAATGCAATTTTAAACCAAGCATGCTCAACATCTTACCT 1132
DB 59148 ACGAAATCTGGGATGCTAATGCAATTTTAAACCAAGCATGCTCAACATCTTACCT 59207
QY 1133 AAAGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1192
DB 59208 AGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 59267
QY 1193 GATGCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1252
DB 59268 AATGAGAAATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 59327
QY 1253 GACCGAGTAATTTCTTACTCTCAATCCGCGGATTTTACATGAGGCTGTTGGCT 1312
DB 59328 GATGAGAGAGATTTTACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 59387
QY 1313 CCTTGATGACAGCATATTTTCTATGCTGCGGATACCTTGTATCTCAACAGTACT 1372
DB 59388 CCTTGATGACAGCATATTTTCTATGCTGCGGATACCTTGTATCTCAACAGTACT 59447
QY 1373 CAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1432
DB 59448 CAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 59507
QY 1433 CTGCGCTGATGACAGATCACAGGAAACCGCGGCTTGGCTGATCACAGATGCTC 1492
DB 59508 TTAGCAATTTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 59567
QY 1493 AATGCGCAATTTACCAAGATCACTGCAATGATGATGATGATGATGATGATGATGAT 1552
DB 59568 AATGCGCAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 59627
QY 1553 CCAATCCAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1612
DB 59628 CCAATCCAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 59687
QY 1613 ACTTGAGTACTTGAATGATTTTGAAGTACCGGATATGATGATGATGATGATGATGAT 1672
DB 59688 ACTTGAGTACTTGAATGATTTTGAAGTATGATGATGATGATGATGATGATGATGAT 59747
QY 1673 TACCGCTTCACTTAT 1691
DB 59748 TACCGCTTCACTTAT 59766
```

RESULT 8
 ID ACA42822 standard; DNA; 852 BP.
 AC ACA42822;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Prokaryotic essential gene #24479.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KM drug design; gene.
 XX
 OS Pasteurella multocida.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 XX
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 PI WPI; 2003-029926/02.
 DR P-PDSB; ABU38952.
 XX
 PT New Antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PS Claim 14; SEQ ID NO 30692; 1766bp; English.
 XX
 XX The invention relates to an isolated nucleic acid comprising any one of
 the 6213 antisense sequences given in the specification where expression
 of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence data for this patent did
 not form part of the claimed specification, but was obtained in
 electronic format directly from WIPO at
 ftp.wipo.int/pub/published/seq_sequences

XX	Sequence	852 BP;	247 A;	201 C;	181 G;	223 T;	0 U;	0 Other;
QY	Query Match	17.4%;	Score 505.4;	DB 7;	Length 852;			
Db	Best Local Similarity	74.6%;	Pred. No. 6.8e-150;					
Matches	635;	Conservative	0;	Mismatches 216;	Indels	0;	Gaps	0;
QY	840	TGAACAGATTTGATCTTTGTGACGGCAATCGTCGATCAAGGTGTTTGGGTTGAAATG	899					
Db	2	TGAACACTATTGAACTTTGGCAAGTATCGTTATGAAAGGTGTTTGGGTTGAAATG	61					
QY	900	AAGCAAGGGCAAGCGTTGTTGACTGTGATTAATGCCGATTGACCTACGATGTGGCA	959					
Db	62	CACGAGCTGGCAACGTGTCTGTGACCGTGAATGATGGGATCTCACCTACGATGTGGCA	121					
QY	960	ACATTCAGTTTCCCTCTAGTACTACAGCAAGATTTTGGAAAGCTGCCGTAGCCGAGT	1019					
Db	122	ATTAACCATTTTCCGTTGATCAACCAAGCAAACTATGGAAGGCCAATTGCGAGAT	181					
QY	1020	TGCTGGCTATATTGTTGTTGATTAATGCGGATTTTCCGCAATTAGTATCCAAA	1079					
Db	182	TTTAGTTATATCCGTGCTATGACATGCGGAGATTTCCGCAAACTCCGCAACAAA	241					
QY	1080	CCTGGGATGCTAATGCCAATTTAAACAGATGGCTCAACATTCCTTACGTTAAAGTG	1139					
Db	242	CCTGGGAGCCAAAGCTATGATAAAACAGTATGGCTCAATTAATCCGACCCGCAAGGCA	301					
QY	1140	AGGATGACATGGAGCGGTGATGATGTCTTACGGGTAGAGCTTTGGGCTTAAGCTGATG	1199					
Db	302	CCGATGATATGGAGCGTGTATTATGGCGTACAAAGGCGGTGATGGCGCAACCGAAGCG	361					
QY	1200	GTCATATTGACCAAGTTGAAAAGATTTGTGATTAATTGAGCCGTGGCGTTGATGACGAG	1259					
Db	362	AAACCGTCGATCAATTTGGCAAAATCGCATTAATCTGGCAAGGCAATTATGACCGTG	421					
QY	1260	GTGAATTCCTTAACCTTCAACATCCGGGTGAATTTCAATGAGGGGTGTTGGCCCTTGCA	1319					
Db	422	GTGAATTCATACCTTTCTTCAACCCGGGTGAATTTGAACTGGGTTGCTTAGGTCCTTGCA	481					
QY	1320	TGTAAGCGCATTTTTCATTGCTGTGGGGATTAACCTGTATCTCAACGTACTACGCTT	1379					
Db	482	TGCACACCCACACTTCCTCATTAACCTGGTGAATACCCCTATCTACCCAGTATACGCGTT	541					
QY	1380	CATGTAATGTGCCCTTGCGGTTGAAATTTCAACATGGTACAGTTATGTTCCTTGCGC	1439					
Db	542	CTTGTAATGTGCCCTTGCGGTTGAAATTTCAACAAATTCAGATGTACCTTTCTTGCT	601					
QY	1440	TGATGGCACATGACAGGAAAAAAGCCGGCTTGCGCTATCACAAATGCTCAATGGC	1499					
Db	602	TGATGGCACAAATTAAGGCAAAAAAGGGGTAAAGCGGTACACAAATTCATCATGAC	661					
QY	1500	ACATTTACCAAGTCACTGCAATTTGATGGCCGATGTCCAGCTAAAGTGAAGCATTC	1559					
Db	662	ACATTTATGAAGATCACTTATCTTCAGAAAAATGTGCAATTTAAAGTGAAGCATTC	721					
QY	1560	CAGGCGCTCAGTCCCATATCAATCCAAAGATTTAAACACTCGAGATTGGAACTTGGG	1619					
Db	722	CGTTCACCAATTACATTAACCCCGAGATTTAAACCTCTGAGATTTTAAAGATTGGG	781					
QY	1620	TCACTTGGATGATTTTGACGTCACCGGATATCAGTTCCAGATCTTATCAATACCGGT	1679					
Db	782	TGACTATGATGATTTTAAAGTCAAGGCGCTATCAATGCCAGATGCCATTAATATCTT	841					
QY	1680	TTTCACTCTAA 1690						
Db	842	TCTCTGTCTAA 852						
XX	ACAA4292 standard; DNA; 852 BP.							
AC	ACAA4292;							

XX 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #15949.
DE Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
OS Haemophilus influenzae.
XX
XX WO20027183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX
XX (BLIT-) ELIIRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HR;
XX WPI; 2003-029926/02.
XX P-P.SDB; ABUS0422.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids, required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 22162; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence data for one of the target
XX prokaryotic essential genes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pat_sequences
XX
XX Sequence 852 BP; 246 A; 152 C; 193 G; 261 T; 0 U; 0 Other;

Query Match 15.9%; Score 462.2; DB 7; Length 852;
Best Local Similarity 71.4%; Pred. No. 4.2e-136;
Matches 608; Conservative 0; Mismatches 243; Indels 0; Gaps 0;

QY 840 TGAACAGTATTAGATCTTTGTGACGACATCGTCATCAAGGTGTTGGGTGAAATG 899
DB 2 TGAAGCAATATCTTGAGCTTTGTGCGGCAATTTGATGAGGGGAATGGTTCATATG 61
QY 900 AACGAAGGGCAAGCGTTGTTGACTGTGATTAATGCGATTGACCTACGATGCGCA 959
DB 62 AACGTACAGGTAAAGCATTCGCTCACGGTCAATTATGACATTTAGATGATGCGCA 121
QY 960 ACATACGTTTCTCTAGTACTACAGCAAGGATTTGGAAAGGCGGTAGCCGAGT 1019
DB 122 ATATCAATTTCCGCGATTAATCAACCGTAAAGTTATGAAAGCGGCGATTCGTAAT 181
QY 1020 TGCTGGCTATATTCGTGTGTTAAGATATGCGGCGATTTTTCGCAATTAGTACCAAA 1079
DB 182 TTTAGGTATATTCGTGATGATGACATCCGCTGATTTCCGCGACTTGACAGAAA 241
QY 1080 CTTGGATGCTAATGCTCAATTTAAACCAAGCATGCTCAACATCTTACCCGTAAGG 1139
DB 242 CTTGGATGCTAATGCTCAATTTAAACCAAGCATGCTGCGTTCGCAATTCGATGAGGCG 301
QY 1140 AGATGACATGAGGACGCGTATGATGATGATGATGATGATGATGATGATGATGATG 1199
DB 302 TTGATGATATGAGGCGCGGTATATGATGATGATGATGATGATGATGATGATGATG 361
QY 1200 GTCATATTGACCGTGAAGAAAGATTTGATGATGATGATGATGATGATGATGATG 1259
DB 362 AAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 421
QY 1260 GTGAATTTCTTAATCTTCAATATCCGCGTGAATTTTCAATGAGGATGTTGCGCGCTTGA 1319
DB 422 GAGGATTTTAACTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 481
QY 1320 TGTACAGCATCAATTTTCTTCTGCTGCGGATGATGATGATGATGATGATGATGATG 1379
DB 482 TGCAATGCGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 541
QY 1380 CATGATATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1439
DB 542 CTTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
QY 1440 TGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1499
DB 602 TTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 661
QY 1500 ACATTTACCAAGTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1559
DB 662 ATATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 721
QY 1560 CAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1619
DB 722 CATTCACCAATTTAGAAATTTATCCAGATTAATTAAGCGTTGAAATTTGAAATTTG 781
QY 1620 TCACTTTGATGATTTTGAAGTCAACCGATATGATGATGATGATGATGATGATGATG 1679
DB 782 TCAAGATGATGATTTTAAAGTGTGCTATCAATCCACAGCAATTAATATATCTT 841
QY 1680 TTTTCAGCTTA 1690
DB 842 TTTTCGCTTA 852
RESULT 10
ACAA4455
ID ACAA4455 standard; DNA; 849 BP.
XX ACAA4455;
XX
XX 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #26112.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;

KM drug design; gene.
 XX
 OS Proteus sp.
 PN WO20027183-A2.
 PD 03-OCT-2002.
 PF 21-MAR-2002; 2002WO-US009107.
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI: 2003-029926/02.
 DR P-PDB; AB040585.
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 14; SEQ ID NO 32325; 1766bp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation; (7) identifying a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 CC
 SO Sequence 849 BP; 263 A; 155 C; 189 G; 242 T; 0 U; 0 Other;

Query Match 13.9%; Score 403.8; DB 7; Length 849;
 Best Local Similarity 67.3%; Pred. No. 1.9e-117;
 Matches 570; Conservative 0; Mismatches 277; Indels 0; Gaps 0;

QY 840 TGAACAGATTTAGATCTTTGTCAGCGCATCGTCGATCAAGCTTTGGTGAATAATG 899
 DB 2 TGAAGAGATCTGCGATCTGTCTCAACGATTTATGATGACGATCAAGCTTAATGATTAATA 61
 QY 900 AACGAACGGGCAAGCGTGTGTTGACTGTGATTAAATGCGGATTGACCTAAGATGTGGCA 959

DB 62 AACGACAGAACCCGTTGTTAAGCTGATTAATGCCGATCTGGATATGATGTTGCCA 121
 QY 960 ACAATCAGTTTCTCTAGTGAATACACGCAAGAGTTTGGAAAGCTGCCGATAGCCGAGT 1019
 DB 122 ATACCAATTTCCATCATTAACAACCGGTAAAGTTTTCACAAAGCCGATTTGCCGAC 181
 QY 1020 TGCTGGGCTATATTCGCTGCTGTTAGATATAGCGGGGTTTTCGCAATTTAGTACCAAAA 1079
 DB 182 TATTAGGACTTACGCTGCTGTTAGATATAGCGGGGTTTTCGCAATTTAGTACCAAAA 241
 QY 1080 CTTGGGATGCTAATGCGCAATTTAAACCAACATGCTCAACAAATCTTACCGTAAAGTG 1139
 DB 242 CGTGAACGGGAATGCTAATGAAATAAGCCGATGCTTAAATTAATCCGATGTAAGGCTG 301
 QY 1140 AGATGACATGGGACGGCTGATGCTTACAGGTAGAGCTTGGGCTAAGCTGATGCTG 1199
 DB 302 AAGATGATATGGGCGGAGTCTATGCTGTTCAAGACGCGCAATGCAACGCGCTGATGGCT 361
 QY 1200 GTCAATTTGACCAAGTTGAAAAAGATTGTTGATGATTTGACCCGCTGCTGATGACCGAG 1259
 DB 362 CGCATTTGATTCATTTACGTAAGTGTGATATTAATTAACGATTTAGACGATGCTG 421
 QY 1260 GTGAATTTCTTAATCTTCAATTCGAGTGGGTTAATTTCAATGCGGCTTTGCCCTTGCA 1319
 DB 422 GTGAGATGACATCTTTTATATATCCGAGAAACAGCGTTAGGCTGTTACGTCATGCA 481
 QY 1320 TGTACAGCCATCAATTTTCAATGCTGGGGGATACCTTGATCTCAACAGTACAGGCTT 1379
 DB 482 TGCTACACATCTTTTCAATGCTGGGGGATACCTTGATCTCAACAGTACAGGCTT 541
 QY 1380 CATGTGATGCTGCTTGGGCTGTAATTTCAACATGCTGACGTTTATGTTCTCTTCCGCG 1439
 DB 542 GCTGTGATGCTGCTTGGGCTGTAATTTCAACATGCTGACGTTTATGTTCTCTTCCGCG 601
 QY 1440 TGAATGACATGACGAGGAAAAAGCGGCTGCTGATTCACAAAGTGTCAATGCGC 1499
 DB 602 TGGTGCTCAATACAGGGGATTAACAGGTTAAGATTCATTAAGATGTTAAATGCTC 661
 QY 1500 ACATTTCCAGATCACTCAATGATGATGCGCATGTCAGCTTAAACGTCAGCCATTC 1559
 DB 662 ACATTTGAGAACCATTTACCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 721
 QY 1560 CAGCGCTCAGTTCCATATCAATCAAGATTTAAACATGCTGACGATTTGAAACCTTGGG 1619
 DB 722 CATTAACCTAAGCTAATATTAATCAAAACATTAAGACATTAAGATGATTAAGAACTTGGG 781
 QY 1620 TCATTTGATGATTTTGAAGTCAACGCGATATGATGCTGATGCTGATGCTGATGCTGATGCTG 1679
 DB 782 TCACGACATGATGATTTAGTTGATTAAGGTTATCAAGTGTCAAGGAGCATTAATACCCGT 841
 QY 1680 TTTTCAGT 1686
 DB 842 TCACGGT 848

RESULT 11
 AA237118
 ID AA237118 standard; DNA; 807 BP.
 AC AA237118;
 XX
 XX 26-JAN-2000 (first entry)
 XX
 XX Nucleotide sequence of H. influenzae H10904 gene.
 XX
 XX Genome; mutagenesis; transposon; isolate; locate; essential gene; ss;
 XX detect; growth; anti-microbial therapy; genomic footprinting.
 XX Haemophilus influenzae.
 XX
 XX W0950402-A1.

WP ACF67367_54 5400001 5510000
WP ACF67367_55 5500001 5610000
WP ACF67367_56 5600001 5648894

Query Match
Best Local Similarity 61.1%; Pred. No. 4.8e-76;
Matches 544; Conservative 0; Mismatches 278; Indels 69; Gaps 3;

14 ATGCTCAGAGGTTATGCAAGTTCCCAATATTGACCCGATGTGTTTGCATCGCCCT 73
94457 ATGAGTACAGACTACCTGCGCATTTCTAATATTGACCCGATATTTCATATAGGTCCA 94398

74 CTAGCGGTGCGGTGATGAGCTTGATTTGGTGGGTTTCCCTTTTGTCTATGTTGGTGG 133
94397 ATAGCCCTTCATTTGATGATGTTTCTATGTTACTTGGTTGGTGGTGGTGGTGGTGA 94338

134 GCCAATGCGCGAGCGGATGCGCGGGGAGTGGTGGAGCGGTGAGCAAGTCTCTGACTTG 193
94337 GCGACACGCGAGGGGACGAAACCAATAGTGGTGGCAAAAACGAAGTAGAAGAACTTG 94278

194 TTAATGCGCGGCTTTTATGATGATGATGCGGCGGAGTGGTGGTATGATCTTTAC 253
94277 CTATATGAGGTTTCTGCGCATTTGTTGGGCGCTGCGGCTACGTCCTGTTTAT 94218

254 AATTGATCTGTTCCCTGCTGACCCCTTTATTTTCAAGTGGAGTGGCGGCAAG 313
94217 AATTGCTGCTTTCTTGCACAAACCCCTTTATTTGTTAAAGTTGGGATGTTGGAGT 94158

314 TCCCTTCAAGCGCGCTTATTTGGGTGATGACCGCATTTCTGTTAGCGCGTAAAC 373
94157 TCTTTCCAGGTTGTTAGTGGGGTATCTGCGCAATGTTGGTGGTGGCGGCAAGC 94098

374 CAACGACCTCTCTTGTGTTGGCCGATTTGTTGCCCCCTTATGTCGATTCGATTTGGG 433
94097 AAACGCAATTTCTTAAGTTGCTGACTTCATTTGCCCTTTAGTCCCTTTTGGTTAGGC 94038

434 ATGAGCGTATCGTAATTTATGAACTGTAAGTGAACCTTTGGGAGCAAGTAAAC---GGATGTG 490
94037 ATGGGAGAGATCGTAATTTATTTAAACGCGAGTTGTGGGAGCGTATCATTTGATACG 93978

491 CCTTGGGCTTTTGTATTCCTA----- 512
93977 CCGTGGGCAATGTTATTTCCAGCTCCGCTGGAAGATATTGCAATGCTGCAACCGAT 93918

513 -----ATGGTGGCCCACTCCGCGCCATCCCTTACAGCTT 547
93917 CCTCAGCTGTAACAGTACTGAGCAATATGGTTCTGCCAAGTCAACCTTCAAGGCTT 93858

548 TATGAATTTGGCTTAAGAGCGTGTCTGTTCTTATTTCTTAATTTGGTTTATTTGGTAA 607
93857 TATGAATTTGGCTTAAGAGCGTGTCTGTTCTTATTTCTTAATTTATTTATTTCCGTAA 93798

608 CCTGCGCCCTAGGAGCGAGTATCCGAGCTTTTATGAGTGAATGAGTCAATTTCCGCTTC 667
93797 CCTGCGCCCTAGGAGCGAGTATCCGAGCTTTTATGAGTGAATGAGTCAATTTCCGAGT 93738

668 CTGTTGGAATGAGTCCGTTAGAGCGAGTCAAGTGTCAAGTGTCTGTTGGTGGCTTCAATTCA 727
93737 ATAGTGAATTTCTTCCGCAACAGATGCAAACTGGGCTTATTTGAGCGGAGTCAAGT--- 93681

728 ATGGGCAATCTCTCTTACCTATGAGTGAATGAGTGAATTTGATGATGTTGGTGGTCT 787
93680 ATGGGCAATTTCTTCTATACCAATGATCTCGCGGGAATTTATATGATGATGAGGCA 93621

788 TACAAGCGGCTTTTATCAAGACCGTGTAGACGCAAAATAGGATGATG 838
93620 TATAACACCAAGTATATAGTACAAAGTAAATGAAGATATCTGG 93570

Db 93620 TATAACACCAAGTATATAGTACAAAGTAAATGAAGATATCTGG 93570

RESULT 13
ACF65386_0
WP Sequence split into 7 fragments LOCUS ACF65386 Accession ACF65386
WP Fragment Name Begin End

WP ACF65386_0 1 110000
WP ACF65386_1 21000
WP ACF65386_2 20001 31000
WP ACF65386_3 30001 41000
WP ACF65386_4 40001 51000
WP ACF65386_5 50001 61000
WP ACF65386_6 60001 700779
ID ACF65386 Standard; DNA; 700779 BP.

ACF65386;
20-NOV-2003 (first entry)

Phototribulus luminescens nucleotide sequence #39.

Antibacterial; fungicide; insecticide; polymorphism; genetic analysis; detection; food; gene expression; plant; animal; microorganism; toxin; antibiotic; biopesticide; virulence factor; disease model; plague; whooping cough; gene; ds.

Phototribulus luminescens.

W0200294867-A2.

28-NOV-2002.

07-FEB-2001; 2001FR-00001659.

07-FEB-2002; 2002WO-IB003040.

(INSP) INST PASTEUR.

(CNRS) CNRS CENT NAT RECH SCI.

Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A, Buchrieser C,

WP1; 2003-148459/14.

Genomic sequence of Phototribulus luminescens and encoded polypeptides, useful e.g. as therapeutic antimicrobials and agricultural pesticides.

Claim 1; SEQ ID NO 39; 1205bp; French.

The invention relates to the isolation of genes and their encoded proteins from Phototribulus luminescens. The isolated sequences are sources of probes and primers for detecting the genome of P. luminescens and related species; to study polymorphisms; for gene analysis and for detection/amplification of the genes. Antibodies (Ab) raised against the polypeptides encoded by the genes are used for detection/identification of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that carry a gene-containing vector are used to select compounds that modulate, regulate, induce or inhibit expression of the genes in plants, animals or microorganisms other than P. luminescens and are able to alter response or sensitivity to toxins and antibiotics produced by P. luminescens. Cells transformed to express the genes are useful for recombinant production of the proteins, particularly toxins and antibiotics. Antibacterial materials useful as insecticides, bactericides and fungicides. The genes, proteins, vectors containing the genes and Ab are also useful therapeutically (to treat microbial infection by bacteria or fungi that are sensitive to P. luminescens-encoded toxins or antibiotics) and as biopesticides. Other uses of the genes and the proteins are as virulence factors and for identifying targets of human diseases for which P. luminescens is a model (particularly plague and whooping cough). This sequence represents one of the isolated P. luminescens genes.

Sequence 700779 BP; 190440 A; 140977 C; 164444 G; 204917 T; 0 U; 1 Other;

Query Match
Best Local Similarity 61.1%; Pred. No. 4.8e-76;
Matches 544; Conservative 0; Mismatches 278; Indels 69; Gaps 3;

14 ATGCTCAGAGGTTATGCAAGTTCCCAATATTGACCCGATGTGTTTGCATCGCCCT 73

Db 104168 ATGAGTAAACGCTACTGCGATTCTCTAATATTGACCCGCTCATTTTTCATAGTGTCA 104227
 QY 74 CTAGCGGTGCGCTGTATGCGCTTGTATGTTGGTGGTTCCTTTTCTATGTTG 133
 Db 104228 ATAGCCCTCATGTGTATGTTTCACTTGTGGTGTGTATTCGCCATGAGGCTA 104287
 QY 134 GCCAATGCGGAGCGGATGCGCGGAGGTGTGACCGGTGAGCAAGTCTCTACTTG 193
 Db 104288 GCCACACGACGAGGACGAAACCAATAGTGTGACAAAAACGAGTATAGAACTTG 104347
 QY 194 TTATGCGCGGCTTTTAAAGTGTAGTATCGGTGCGGAGTTGTATGTATCTTAC 253
 Db 104348 CTTATGACAGTTTCTGCGCATTTGTGTGGCGCTGTGCGCTACCTCTGTTTAT 104407
 QY 254 AATTGTATGTTCTCTGCTGACCCCTTTTATTTATCAAGTGTGACGTGGGCAATG 313
 Db 104408 AATTCCCTGCTTCTTCGACCAACCCCTTTATTTTAAAGTTGGATGTGGATG 104467
 QY 314 TCCCTCAGCGCGCTTATTTGGTGTATGATCACCGCATGTCTGTATGCGGTAATAAC 373
 Db 104468 TCTTTCAGCGGTGTATGAGGGGTTATCTGCGCATGTGGTGTGGCGGAGAACG 104527
 QY 374 CAACGCACTTCTTTGCTGTGGCGCATTTGTGCCCCCTTGTAGTGTGCTTGGG 433
 Db 104528 AAACGCAATTTCTTACAGTGTCTGACTTCAATGCCCCGTATGCTCTTGTGTTAG 104587
 QY 434 ATGGACGATCGGTACTTATGATATGATGAACTTTGGGACGAGTAAC--GGATGTG 490
 Db 104588 ATGGGAGAAATCGGTAAATTTTATTAACGAGATTGTGGGACGTGTCAATGTGATCG 104647
 QY 491 CTTGGGCTTTTATTCCTTA----- 512
 Db 104648 CCGTGGCGCATGTTATTTCCCAAGCTCCCGTGTGAGATATGCAATGCTGCAACCAT 104707
 QY 513 -----ATGGGCGCCACTGCGCGCATCTTCAACAGCT 547
 Db 104708 CCGTCACTGTATCAAGTACTGAGCAATAGTGTCTTCCACGTACCTTCAACAGCT 104767
 QY 548 TATGAATGCGCTTAAAGCGGTGTGTTCTTATTTCTTATTTGTTATGTTAA 607
 Db 104768 TATGAATGCGCTTAAAGCGGTGTGTTCTTATTTGATTTATTAATATTCGTTAA 104827
 QY 608 CCGTCCCGCTAGGACGCGTATCCGAGCTGTTTATGCTGGAATACGTTACCTTCCGCTTC 667
 Db 104828 CCGTCCCGCAATGGGCAAGTGTTCGGTATTCTTATTTAGGCTACGATATTTCCGAGTG 104887
 QY 668 CTTGTGAATACGTCCGTGAGCGAGATGCTCAATGGGTCTGTTGGGCTTCAATTTCA 727
 Db 104888 ATAGTGAATGTTCTCGCCACAGATGACAACTGGCTTATTCGACGGGATCAAGT--- 104944
 QY 728 ATGGGGCAATCTCTCTTACTATGTTGATCATGCTATTTTGAATGTTGGTGT 787
 Db 104945 ATGGGGCAATTTCTTCTATACCAATGATCTGGCGGAAATATTAATGATGATGGGCA 105004
 QY 788 TACAAGCGCGTTTGTATCAAGACCTGTAGACGAAATAGGATAGTAG 838
 Db 105005 TATTAACACCAAGTATATAGTACAAAGGATTAATGAACGATATCTGG 105055

RESULT 14
 ACF5386_1
 Continuation (2 of 7) of ACF5386 from base 100001 (Photobhabdus luminescens nucleotide)

WP Sequence split into 7 fragments LOCUS ACF5386 Accession ACF5386

WP	Fragment Name	Begin	End
WP	ACF5386_0	1	110000
WP	ACF5386_1	100001	210000
WP	ACF5386_2	200001	310000
WP	ACF5386_3	300001	410000
WP	ACF5386_4	400001	510000
WP	ACF5386_5	500001	610000
WP	ACF5386_6	600001	700779

Query Match 9.6%; Score 278.2; DB 7; Length 110000;

Best Local Similarity 61.1%; Pred. No. 4,8e-76;
 Matches 544; Conservative 0; Mismatches 278; Indels 69; Gaps 3;

QY 14 ATGCTCAGAGGTATATCGACATTTCCCAATATGACCCCGTATGTTTGCATGCGCCCT 73
 Db 4168 ATGAGTAAACGCTACTGCGATTCTCTAATATTGACCCGCTCATTTTTCATAGTGTCA 4227
 QY 74 CTAGCGGTGCGCTGTATGCGCTTGTATGTTGGTGGTTCCTTTTCTATGTTG 133
 Db 4228 ATAGCCCTCATGTGTATGTTTCACTTGTGGTGTGTATTCGCCATGAGGCTA 4287
 QY 134 GCCAATGCGGAGCGGATGCGCGGAGGTGTGACCGGTGAGCAAGTCTCTACTTG 193
 Db 4288 GCCACACGACGAGGACGAAACCAATAGTGTGACAAAAACGAGTATAGAACTTG 4347
 QY 194 TTATGCGCGGCTTTTAAAGTGTAGTATCGGTGCGGAGTTGTATGTATCTTAC 253
 Db 4348 CTTATGACAGTTTCTGCGCATTTGTGTGGCGCTGTGCGCTACCTCTGTTTAT 4407
 QY 254 AATTGTATGTTCTCTGCTGACCCCTTTTATTTATCAAGTGTGACGTGGGCAATG 313
 Db 4408 AATTCCCTGCTTCTTCGACCAACCCCTTTATTTTAAAGTTGGATGTGGATG 4467
 QY 314 TCCCTCAGCGCGCTTATTTGGTGTATGATCACCGCATGTCTGTATGCGGTAATAAC 373
 Db 4468 TCTTTCAGCGGTGTATGAGGGGTTATCTGCGCATGTGGTGTGGCGGAGAACG 4527
 QY 374 CAACGCACTTCTTTGCTGTGGCGCATTTGTGCCCCCTTGTAGTGTGCTTGGG 433
 Db 4528 AAACGCAATTTCTTACAGTGTCTGACTTCAATGCCCCGTATGCTCTTGTGTTAG 4587
 QY 434 ATGGACGATCGGTACTTATGATATGATGAACTTTGGGACGAGTAAC--GGATGTG 490
 Db 4588 ATGGGAGAAATCGGTAAATTTTATTAACGAGATTGTGGGACGTGTCAATGTGATCG 4647
 QY 491 CTTGGGCTTTTATTCCTTA----- 512
 Db 4648 CCGTGGCGCATGTTATTTCCCAAGCTCCCGTGTGAGATATGCAATGCTGCAACCAT 4707
 QY 513 -----ATGGGCGCCACTGCGCGCATCTTCAACAGCT 547
 Db 4708 CCGTCACTGTATCAAGTACTGAGCAATAGTGTCTTCCACGTACCTTCAACAGCT 4767
 QY 548 TATGAATGCGCTTAAAGCGGTGTGTTCTTATTTCTTATTTGTTATGTTAA 607
 Db 4768 TATGAATGCGCTTAAAGCGGTGTGTTCTTATTTGATTTATTAATATTCGTTAA 4827
 QY 608 CCGTCCCGCTAGGACGCGTATCCGAGCTGTTTATGCTGGAATACGTTACCTTCCGCTTC 667
 Db 4828 CCGTCCCGCAATGGGCAAGTGTTCGGTATTCTTATTTAGGCTACGATATTTCCGAGTG 4887
 QY 668 CTTGTGAATACGTCCGTGAGCGAGATGCTCAATGGGTCTGTTGGGCTTCAATTTCA 727
 Db 4888 ATAGTGAATGTTCTCGCCACAGATGACAACTGGCTTATTCGACGGGATCAAGT--- 4944
 QY 728 ATGGGGCAATCTCTCTTACTATGTTGATCATGCTATTTTGAATGTTGGTGT 787
 Db 4945 ATGGGGCAATTTCTTCTATACCAATGATCTGGCGGAAATATTAATGATGATGGGCA 5004
 QY 788 TACAAGCGCGTTTGTATCAAGACCGTGTAGACGAAATAGGATAGTAG 838
 Db 5005 TATTAACACCAAGTATATAGTACAAAGGATTAATGAACGATATCTGG 5055

RESULT 15

ACF5784
 ID ACF5784 standard; DNA, 876 BP.

ACF5784;

20-NOV-2003 (first entry)

Photobhabdus luminescens nucleotide sequence #8251.

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KM antibiotic; biopesticide; virulence factor; disease model; plague;
 KM whooping cough; gene; ds.
 XX Photorhabdus luminescens.
 OS
 PN WO200294867-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 07-FEB-2002; 2002WO-IB003040.
 XX
 PR 07-FEB-2001; 2001FR-00001659.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Duchaud E, Tsourit S, Glaeser P, Frangaul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX
 DR WPI; 2003-148459/14.
 XX
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 PS Claim 2; SEQ ID NO 8251; 1205bp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes
 CC
 SQ Sequence 876 BP; 200 A; 177 C; 221 G; 278 T; 0 U; 0 Other;

Query: Match 9.5%; Score 277.8; DB 7; Length 876;
 Best Local Similarity 62.2%; Pred. No. 3.2e-77;
 Matches 528; Conservative 0; Mismatches 252; Indels 69; Gaps 3;

QY 14 ATGCTCAGGGTTATCTGCATTTCCCAATATGACCCCGATTTGTTGATCGGCCCT 73
 Db 1 ATGATTAACAGCTACCTTGCAATTCCTTAATATGACCGGATTTTTCATAGATGCA 60
 QY 74 CTAGGGATGCGCTGATATGCTGATGATTTGGTGGGTTCTTTTGGTATGATGATG 133
 Db 61 ATAGCCCTCCATTTGATATGTTTCAATGATTTGGTGGGTTTATTCGCAATGCTA 120
 QY 134 GCCATGCCCGACGGGATCGCGCGGAGCTGTGACGCTGAGCAAGTCTGACTTG 193
 Db 121 GCGACACGACGCGGACGCAAAACCAATAGTGTGACAAAAAAGAAAGTAGAAGACTTG 180
 QY 194 TTATTCGCGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 253
 Db 181 CTTATGAGGTTTTCGCGGATTTTGTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 240

QY 254 AATTTGATCTGTTCTTCTGACACCTCTTATTTATTCAAAGTGTGACATGCGGCATG 313
 Db 241 AATTTCCCTGCTTCTGACACACCTCTTATTTATTTAAAGTTGGAGTGGGATG 300
 QY 314 TCTTTCCACGCGGCTTATTTGGTGTATGATGATGATGATGATGATGATGATGATG 373
 Db 301 TCTTTCCACGCGGCTTATTTGGTGTATGATGATGATGATGATGATGATGATGATG 360
 QY 374 CAAGCACCTTTCTTGTGTGTGATGATGATGATGATGATGATGATGATGATGATG 433
 Db 361 AAAGCCATTTCTTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 420
 QY 434 ATGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 490
 Db 421 ATGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 QY 491 CTTTGGGCTTTTGTATTCCTA----- 512
 Db 481 CCGTGGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 QY 513 -----ATGTTGGCCACTGCGCGCCATCTTACAGCTT 547
 Db 541 CCTCACTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 QY 548 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 607
 Db 601 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 QY 608 CTTTGGGCTTTTGTATTCCTA----- 667
 Db 661 CTTTGGGCTTTTGTATTCCTA----- 720
 QY 668 CTTTGGGCTTTTGTATTCCTA----- 727
 Db 721 ATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 777
 QY 728 ATGGGCAAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 787
 Db 778 ATGGGCAAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 837
 QY 788 TACAAGGC 796
 Db 838 TATTAACAC 846

Search completed: July 31, 2004, 13:34:23
 Job time : 770 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 12:39:50 ; Search time 150 Seconds
(without alignments)
10762.348 Million cell updates/sec

Title: US-09-700-712a-1

Perfect score: 2909
Sequence: 1 ggaaggttctatgctc.....cagcaaatcgcgctgcag 2909

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A.COMB.seg: *
2: /cgn2_6/prodata/2/ina/5B.COMB.seg: *
3: /cgn2_6/prodata/2/ina/6A.COMB.seg: *
4: /cgn2_6/prodata/2/ina/6B.COMB.seg: *
5: /cgn2_6/prodata/2/ina/PCTUS.COMB.seg: *
6: /cgn2_6/prodata/2/ina/backlist1.seg: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	744.2	25.6	1830121	4	US-09-557-884-1 Sequence 1, Appli
2	744.2	25.6	1830121	4	US-09-643-990A-1 Sequence 1, Appli
3	410.6	14.1	873	4	US-09-543-681A-2282 Sequence 2282, Ap
4	292.2	10.0	807	3	US-09-277-565-17 Sequence 17, Appl
5	282.4	9.7	810	4	US-09-252-991A-15221 Sequence 15221, A
6	282.4	9.7	954	4	US-09-252-991A-15232 Sequence 15232, A
7	282.4	9.7	1434	4	US-09-252-991A-15227 Sequence 15227, A
8	265.8	9.1	831	4	US-09-243-681A-2365 Sequence 2365, Ap
9	265.8	8.8	879	4	US-09-489-039A-2965 Sequence 2965, Ap
10	208.6	7.2	834	4	US-09-328-352-3810 Sequence 3810, Ap
11	152.2	5.2	640681	4	US-09-790-988-1 Sequence 1, Appli
12	143.8	4.9	891	4	US-09-540-236-663 Sequence 663, Appl
13	143.8	4.9	96109	4	US-09-596-002-35 Sequence 35, Appl
14	142.2	4.0	435	4	US-09-252-991A-15244 Sequence 15244, A
15	75.4	2.6	2288	4	US-08-535-433-1 Sequence 1, Appli
16	75.4	2.6	2288	4	US-09-553-132-1 Sequence 1, Appli
17	75.4	2.6	4137	4	US-09-499-964-2 Sequence 2, Appli
18	71.6	2.5	2058	4	US-09-252-991A-15243 Sequence 15243, A
19	64.8	2.2	942	4	US-09-347-878-29 Sequence 29, Appli
20	64.8	2.2	1536	3	US-09-089-195-1 Sequence 1, Appli
21	64.8	2.2	1536	4	US-09-367-007C-38 Sequence 38, Appli
22	64.8	2.2	1536	4	US-09-362-665-6 Sequence 6, Appli
23	64.8	2.2	1536	4	US-09-963-333-6 Sequence 6, Appli
24	61.2	2.1	978	4	US-09-134-000C-941 Sequence 941, App
25	57	2.0	975	4	US-09-107-532A-2491 Sequence 2491, Ap
26	54.2	1.9	7685	4	US-09-221-017B-1092 Sequence 1092, Ap
27	49	1.7	969	4	US-09-134-001C-357 Sequence 357, App

28	48.2	1.7	870	4	US-09-252-991A-15224 Sequence 15224, A
29	47.4	1.6	801	4	US-09-489-039A-2884 Sequence 2884, Ap
30	46.6	1.6	867	4	US-09-107-532A-2355 Sequence 2355, Ap
31	46.6	1.6	1014	3	US-09-230-637-1 Sequence 1, Appli
32	46.6	1.6	35100	2	US-08-770-379-17 Sequence 17, Appli
33	46.6	1.6	35100	3	US-08-757-669A-17 Sequence 17, Appli
34	46.6	1.6	35100	4	US-09-230-371A-17 Sequence 17, Appli
35	45.8	1.6	5924	4	US-08-956-171E-130 Sequence 130, App
36	45.4	1.6	2335	4	US-08-961-527-275 Sequence 275, App
37	44.8	1.5	852	4	US-09-134-001C-2654 Sequence 2654, App
38	44.8	1.5	870	4	US-09-134-000C-1317 Sequence 1317, Ap
39	42.6	1.5	621	4	US-08-935-433-3 Sequence 3, Appli
40	42.6	1.5	621	4	US-09-583-132-3 Sequence 3, Appli
41	42.4	1.5	1839	4	US-09-203-895-2 Sequence 2, Appli
42	40.6	1.4	7218	1	US-08-232-463-14 Sequence 14, Appli
43	40.6	1.4	1568	4	US-08-936-165A-95 Sequence 95, Appli
44	39.8	1.4	4214	4	US-09-221-017B-293 Sequence 293, App
45	38.6	1.3	7218	1	US-08-232-463-14 Sequence 14, Appli

ALIGNMENTS

RESULT 1
US-09-557-884-1
Sequence 1, Application US/09557884
Patent No. 650681
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESSES:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557, 884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1
Query Match
Best Local Similarity 66.7%; Pred. No. 1.7e-245;
Matches 1134; Conservative 0; Mismatches 538; Indels 27; Gaps 4;
2 AGAAGTTTGTATGCTCAGGCTTATGCTGATTTCCCAATATTGACCCCGTATTGTTT 61

Db 958086 ATAGAGAAATTTAGATTTAAATTTATCTTCTCCCACTTTGATCCGATTTT 958145
 Qy 62 TCATCGGC-----CCTTAGCGGTCGCTGCTGATGCTTATGATTTGGTGGCTTC 115
 Db 958146 ACGCTTGCCGATGATATATATCGGCTTACGTTGATGCTTGAATACCTTTTAACTTT 958205
 Qy 116 CTTTTCGCTATGCTTGGCCCAATCGCCGAGCGGATCGCGGCGATGCTTGAAGCGCT 175
 Db 958206 GTTTTTCACAGTGGCTTGGCTTCCCGCTGATACGCCAAATGCGGTTGACAGTA 958265
 Qy 176 GAGCAAGCTCTGATCTTATTTGCGCGCTTTTATGCTGATGATCGTGGCGGATTT 235
 Db 958266 GATCAAGTATGATGCTTACTTTTCAACGCTTTTATGGGCGTGTATTTGCGGACGCTGT 958325
 Qy 236 GCTTATGCTATCTTCAATTTTGAATGCTGCTGCTGATGATCGCTGCTTATTTATCA 295
 Db 958326 GCGCATATATTTTCTATATCTCGATCATTTCTTACAGAACCACTTTATTTATCCG 958385
 Qy 296 GCTGGAAGTGGCGGATGCTTCCACAGCGGCTTATGGGCTGATCAACCGCATGTTG 355
 Db 958386 GTTTGGAGAGTGGAATGCTGCTTCCACGCTGCTTATGCTTATGCTTATGCTATGAT 958445
 Qy 356 TGTATCGCGCTTAAAAACCAAGCACTTCTTGTGCTGCGATTTGTTGCCCTTAA 415
 Db 958446 TGGACATCTTATCTCAAAAACGTAATTTTGGCAACGCTATTTTGTGCGCTTTG 958505
 Qy 416 GTGCGATTCGTTTGGGATGGAGCGATCGTATCTTATGATAGTGAACCTTTGGGA 475
 Db 958506 ATTCGCTTTGGTTTAAAGTTTAAAGGAAATGCTTATCTTATGATGATGATGAGGA 958565
 Qy 476 CGAGTAACGATGCTGCTTGGGCTTTTGTATTCCTTATGATGATGCTGCTGCTGCT 532
 Db 958566 CGCGAAACGATGCTTGGGCAATGATTTCCCGATGATCTCTTTTACGCTGCTG 958625
 Qy 533 CATCTTCACAGCTTATGATGCTGCTTGAAGCGCTGCTGCTTCTTATCTTAA 592
 Db 958626 CATCATCAACAATTTGAAGCTTTTGAAGCGCTTGTGTTTGAAGCTTGAAT 958685
 Qy 593 TGTATTATGTAACCTGCTCGCTAGGCAAGGATCGGATCGGACTTTTATGCTGATAC 652
 Db 958686 ATTTTATTAATAAACACGCTCCATGCTTCTGTTGCAAGTTTATCTTATGCTTAT 958745
 Qy 653 GGTACATTCGCTCTTGTGGAATAGTCGCGTAGGCAAGTCTAGTGGCTGTT 712
 Db 958746 GGGCTCTTCGCTTTATTTGGAATATGCTGTAACCTGAG-----TTGAAATTTTC 958799
 Qy 713 GGTGCTTCATTTCAATGCGGCAATCTCTCTTACCTATGCTGATCGGATTTTG 772
 Db 958800 TTGGAATATTTACACAGGCGCAAGCCCTTGTGCGGATGATTAATGCTGCTTTC 958859
 Qy 773 ATGATGCTTGTGCTTACAGCGGCTTGTGATCAAGCCGTGAGCAAGCAAAATGAGGT 832
 Db 958860 ATTTATGCTTGGGCTTTATTCACGC-----AAAGTGGGCTCATTAATAATGAG 958907
 Qy 833 AGTAGGTAAGAAAGATTTAGATCTTTGTACAGGCAATCGTCAAGGTGTTGGGT 892
 Db 958908 ATTTTATGAAGCAATCTTGAAGCTTTGTGCGGCAATGTTAGTGAAGGAGATGGGT 958967
 Qy 893 GAAATGACGAAGGCGAAGCGCTTGTGATCTGTGATTAATGCCATTTGACTTACGAT 952
 Db 958968 GCTATGAACGTAACGATGATGCTTCAACGCTCATTAATGAGATTTAGAAATGAT 959027
 Qy 953 GTGGGCAACATGATTTCTCTGATGATCAACGCAAGATTTTGAAGCTGCGCTA 1012
 Db 959028 GTGGGCAATATCAATTTCCGCTATTAATCCGTTAAAGTATTTGAAGAGCGGAT 959087
 Qy 1013 GCCGAGTTGCTGGCTATTTCTGCTGATGATATGCGGAGATTTTCCCAATTAAGT 1072
 Db 959088 GCTGAATTTTAAAGTTTATCTGATGATGACATGCGCTGATTTTCCGCACTTGGC 959147
 Qy 1073 ACCAAACCTGGAGTGAATGCAATTTAAACCAAGATGGCTCAACATCTTACCGT 1132

Db 959148 ACGAAACCTGGAGTGTAAATGCGAATGAATAATGACCTTGGCTTGCAAAATCCGATCGT 959207
 Qy 1133 AAAGTGAGATGATGATGAGACGCTGATGATGATGATGATGATGATGATGATGATGAT 1192
 Db 959208 AAGAGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959267
 Qy 1193 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1252
 Db 959268 AATGAGAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959327
 Qy 1253 GACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1312
 Db 959328 GATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959387
 Qy 1313 CTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1372
 Db 959388 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959447
 Qy 1373 CAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1432
 Db 959448 CAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959507
 Qy 1433 CTTGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1492
 Db 959508 TTAGACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959567
 Qy 1493 AATGCGCATTTTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1552
 Db 959568 AATGCGCATTTTCAACGATGATGATGATGATGATGATGATGATGATGATGATGAT 959627
 Qy 1553 CCAATCCAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1612
 Db 959628 CTTTCCCATTAACAAATTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 959687
 Qy 1613 ACTTGGTACTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1672
 Db 959688 ACTTGGTACTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959747
 Qy 1673 TACCGCTTTCAGTCTAAT 1691
 Db 959748 TATCCTTTTCGCTAAT 959766

RESULT 2
 US-09-643-990A-1
 ; Sequence 1, Application US/09643990A
 ; Patent No. 6528289
 ; GENERAL INFORMATION:
 ; APPLICANT: Robert D. Fleischmann
 ; Mark D. Adams
 ; Owen White
 ; Hamilton O. Smith
 ; J. Craig Venter
 ; TITLE OF INVENTION: The Nucleotide sequence of
 ; the Hemophilus influenzae Rd Genome, Fragments
 ; thereof, and Uses Thereof
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville,
 ; STATE: MD
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3 1/2 inch diskette
 ; COMPUTER: Dell Pentium
 ; OPERATING SYSTEM: MS DOS v6.22
 ; SOFTWARE: ASCII text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/643, 990A
 ; FILING DATE: 23-Aug-2000
 ; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: P8186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 25.6%; Score 744.2; DB 4; Length 1830121.
Best Local Similarity 66.7%; Pred. No. 1.7e-245;
Matches 1134; Conservative 0; Mismatches 538; Indels 27; Gaps 4;

QY 2 AGAAGTTTGTATGCTCAGGTTATCTGCACTTCCCAATATTGACCCCGATTGTTT 61
DB 958086 ATAGGAAATATGATTCATTAATTTACTCTTCCCACTTGATCCGATATTTT 958145

QY 62 TCGATCGGC-----CCTTAGCGGTCGCTGTTAGCTTATGTTATTTGGTGGTTTC 115
DB 958146 ACCCTTGGCCATAGTATATACGCTTACGTTGATGCTGATGTAACCTTTTAGCTTT 958205

QY 116 CTTTGTCTATGTTGTTGGCCATGCGCCGAGCGGATGCGCGGCGAGTGGTGGACGCT 175
DB 958206 GTTTTGGACGTTGGCTGCTGGCTGCGCCGCTTAATCGCCCAATACGCGTTGGACAGTA 958265

QY 176 GACCAAGCTCTGACCTTTTATTCGCGCGCTTTTATGTTAGTGTATGATCGTGGCCAGTT 235
DB 958266 GATCAAGTTGATGCTTACTCTTCAACGCTTTTATGCGGCTGTTTATGCGCGACGTTGT 958325

QY 236 GGTATGATCTCTTCAATTTTGAATCTGTCTGCTGACCCCTTATTTATTTCAAA 295
DB 958326 GGGATGATTTTCTTATTTCTCGAATCTTTTCAAGAACCACTTATTTATTCGCG 958385

QY 296 GTGTGACCTGCGGCAATGTCCTTCAACGCGCGCTTATGCGGTGATCAACGCCATGTTTC 355
DB 958386 GTTGGGAAGTGAATGTCGTCACGCGTGGCTTATGTTGATATTTGTTGCTATGATT 958445

QY 356 TGGTATGCGGTAACCAACGCACTTCTTTGCTGTGCGCGATTTTGTGCCCCCTTTA 415
DB 958446 TGGACATCTTATCTCAAAAAGCTATTTTGGCAAAACGCGATTTTGTGCGCCCTTTG 958505

QY 416 GTGCCATTCGTTTGGGGAATGGACGTATCGTAATCTTATGAATAGTAACTTTGGGA 475
DB 958506 ATTCGCTTTGGTTAGCTTTAGGCAAAATGTAATTTGATTAATCTTAACTTAATGGGA 958565

QY 476 CGAGTAAAGGATGCTTGGGCTTTTGTATTCCTTAATGTTGGCC--CACTGCGCGCG 532
DB 958566 CGGAAACGAATGCTTGGGCAATGATTTTCCCAATGATCCTTTTACCTGCTGCT 958625

QY 533 CATCTTCACAGCTTATGATATCGCTTAGAAGGGGTGCTTCTGTTCTTTATCTTAAT 592
DB 958626 CATCTTCACAGCTTATGATAGCTTTTGAAGGCTGCTGTTGTTTACGATCTGAAT 958685

QY 593 TGGTTTATGTTAACTCTGCTCGTAGGCAAGTATCGGACTGTTTATTAAGTGAATAC 652
DB 958686 ATTTTATTAATAAACACAGCTCAATGCGCTCTGTTGACAGTTTATTTCTTAATGTTAT 958745

QY 653 GGTACATTCGCTTCTGTTGAATACGTCGTTGAGCCGATGCTCAGTTGGTCTGTTT 712
DB 958746 GCGCTTCTCGTTTATTTGTTGAATATGTCGTTGAACCTGAAG-----TTGAAATATTC 958799

QY 713 GGTGCTTCATTTCAATGCGGCAATCTCTTCCTTACTTATGTTGATCATGCTATTTTG 772
DB 958800 TTTGGATTTATTAACAGAGGCGACCCCTTGTCTTCCGATGATATTTGTTGCTTTTC 958859

QY 772 ATGATGTTTGTCTTCAAGCGCGGTTTGTATGAAGCCGTTGACGCAAAATAGGAT 832
DB 958860 ATTATGCTTGGCTTATTTACAGC-----AAAATGCGGCTATTAATAATAGAG 958907

QY 833 ACTTAGTGAACAGTATTTAGATCTTTTGTACGCGCATGTCGATCAAGAGTGTGGTT 892
DB 958908 ATTTTATGAAGCAATATCTTGAGCTTTGTGCGCGCATTTGTAGGAAGGGAATGGTT 958967

QY 893 GAAATGGAACGAGGCGAAGGCTGTTTGTACTGATGATTAATGCGGATTTGACCTACAT 952
DB 958968 GCTTAATGAACGTAAGTAAGCATTTGCTCAAGCATTAATGAAGATTAATGAT 959027

QY 953 GTGGCAACAAATCACTTCTCTAGTACTACAGCAAGATTTTGGAAAGCTGCCGTA 1012
DB 959028 GTGGCAATTAATCAATTTCCGCTGATTACTACCGGTAAAGTTTGAAGCGCGCAT 959087

QY 1013 GCCGATGCTCGGCTATATTCGTGTTACGATTAATGCGCGGATTTTGGCAATTAGCT 1072
DB 959088 GCTGAATTTTATGTTATATTCGTGATATACATGCGGCTGATTTCCGCGCACTTGGC 959147

QY 1073 ACCAAACCTGGAATGCTAATGCCAATTTAAACCAAGCATGCTCAACATCTTACCGT 1132
DB 959148 ACGAAATCTGGAATGCTAATGCCAATTAATGAATGCAAGCTTGGCTTGCATTCGATCT 959207

QY 1133 AAGGTGAGATGACATGGAACGCGTATGCTGTTACAGGTTAGAGCTTGGGCTAAGCT 1192
DB 959208 AAGGCGTTGATGATGATGAGGCGGCTATGATGCTGCAAGCAAGCATGCGCTAAGCT 959267

QY 1193 GATGTTGCTCATTTTGAACGATTAAGAAAGTTGTATGATTTGAAGCCGTTGGCTGAT 1252
DB 959268 AATGGAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 959327

QY 1253 GACCGAGTGAATTTCTTAATCTTACAAATCGGATGAAATTTCAATGAGGTTGTTGCGC 1312
DB 959328 GATAGAGAGAGATTTTAACTTTTAAACCTGCGGATTTGATCTTGGTTGCTTGT 959387

QY 1313 CCTTGACATGACAGCATATTTTTCATGCTGCGGAGTATCTTGTATCTACAGTACT 1372
DB 959388 CTTGTATGCAATGCAATCTTCTTCTGTTGGCGATCTTAACTTAACTTAACTGAT 959447

QY 1373 CAGGCTCATGATGATGCGCTTGGGTTGAATTTCAACATGAGGAGTTTATGTTGCTC 1432
DB 959448 CAGGCTTCTGATGATGTTCCGCTTGAATGATTTCAATCAATTCAGGTTTACCTTC 959507

QY 1433 CTGCGCTGATGGAACGATGACAGGAAACCGCGCTTGGCGTATCAACAGATGCTC 1492
DB 959508 TTACACATTTAGGCTGATGATGACCGGCAAAAAGCCGCAAGCATATCAATAATGTG 959567

QY 1493 AATGCGACATTTTCAAGATCAACATGCAATTTGATGCGATGCGCATTAACAGGAG 1552
DB 959568 AATGCGATATTTTGAAGATGACCTTGAATGCTGATGCTGATCACTTAACAGGAG 959627

QY 1553 CCAATCCAGCGCTCACTTCAATATCAATCAAAAGATTAACCACTGACGAGATTTGAA 1612
DB 959628 CATTCCATTAACCAAAATTAAGAAATTAATCAAGATTAATAAAGCGCTTAAGATTGAA 959687

QY 1613 ACTTGGGCTCACTTGAATGATTTTGAAGTCAACGCGATATAGTTTCCAGATCTTATCAA 1672
DB 959688 ACTTGGGCTCACTTGAATGATTTTGAAGTCAACGCGATATAGTTTCCAGATCTTATCAA 959747

QY 1673 TACCGTTTCACTAAT 1691
DB 959748 TATCTTTTCCGCTAAT 959766

RESULT 3
US-09-543-681A-2282
; Sequence 2282, Application US/09543681A
; Patent No. 6605709

GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709,1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO: 2282
LENGTH: 873
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-2282

Query Match 14.1%; Score 410.8; DB 4; Length 873;
Best Local Similarity 67.0%; Pred. No. 5,2e-132;
Matches 583; Conservative 0; Mismatches 287; Indels 0; Gaps 0;

QY 820 AGCAAAATAGGTTAGTTAGTGAACAGTATTTAGATCTTTGTACGGCATGCTGATCA 879
DB 3 AAAAAAAAAAGAGACTACATGAGAGAGATCTGCGATTGTGCAACGATTAATCGATGA 62

QY 880 AGCTGTTGGGTGAAAATGAAAGAAAGGCAAGCGTTGTTGACTGATTAATGCCGA 939
DB 63 AGGTGAATGATGATGATTAACGAGACAGAACCCGTTGTTAACGGTGAATAAGCCGA 122

QY 940 TTGACCTACGATGAGGCAACATCATGTTCTCTAGTACTACGCAAGATTTTG 999
DB 123 TCTGGAATGATGATGTTCCCATTAACCATTTCCATCATCAACACCGTAAAGTTTAA 182

QY 1000 GAAGCTGCGGTACCCGATGCTGCTATATTCGGGTATGATATGCGGCGGATTT 1059
DB 183 CAAGCGCGCATGCGGAATCTAATAGGTACTTAACGGTGTATGATAATGCAACAATTT 242

QY 1060 TCGCCATTAAGTACCAAACTGGATGCTAATGCCAATTTAAACCAAGCATGCTCAA 1119
DB 243 TCGGCTATGTCGTGTAATGCGGAACGCAATGCTAATGAATAATGCGCATGTTAA 302

QY 1120 CAATCTTACCGTAAAGTGAAGTACATGAGGACGGGTATGAGTTCAGGGTACAG 1179
DB 303 TAATCCCATGCTGAAGGTAAAGTGAATGAGGCGGATGATGAGTTCAGGACGCA 362

QY 1180 TTGGGCTAAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1239
DB 363 ATGCAACGCCCTTAATGCTGCACTTTGATCAATTAAGTGGTGAATACTTAAC 422

QY 1240 CCGTGGCGTGAATGACGAGGTGAATTTTAACTTCAATTCGGGTGAATTTACAT 1299
DB 423 TAACGATTAAGCATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 482

QY 1300 GGGGTGTTGGCCCTTGCATGATGATGATGATGATGATGATGATGATGATGAT 1359
DB 483 AGGGTGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 542

QY 1360 TCTCAACAGTACTAGCGTTCATGATGATGATGATGATGATGATGATGATGATGAT 1419
DB 543 TTAAACCTCATATCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 602

QY 1420 GGTATATGTTCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1479
DB 603 ATGCTGTATATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 662

QY 1480 TCACAGATGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1539
DB 663 CCAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722

QY 1540 GCTAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1599
DB 723 ACTAAAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 782

QY 1600 GAGGATTTGAACTTGGTCACTTTGATGATGATGATGATGATGATGATGATGAT 1659

DB 783 AGATGATATGAACCTTGGTCAAGAGATGATGATGATGATGATGATGATGATGAT 842

QY 1660 CGATCTATTCAATACCGCTTTTCACTTA 1689

DB 843 CGAGGCGATTAATACCGCTTCAAGGTA 872

RESULT 4
US-09-277-565-17
Sequence 17, Application US/09277565
Patent No. 6207384
GENERAL INFORMATION:
APPLICANT: Mekalanos, John J.
APPLICANT: Akertley, Brian J.
APPLICANT: Rubin, Eric J.
APPLICANT: Camilli, Andrew
TITLE OF INVENTION: SYSTEMATIC IDENTIFICATION OF ESSENTIAL
FILE REFERENCE: 00742/052002
CURRENT APPLICATION NUMBER: US/09/277,565
CURRENT FILING DATE: 1999-03-26
EARLIER APPLICATION NUMBER: 60/079,770
NUMBER OF SEQ ID NOS: 29
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 17
LENGTH: 807
TYPE: DNA
ORGANISM: Haemophilus influenzae
US-09-277-565-17

Query Match 10.0%; Score 292.2; DB 3; Length 807;
Best Local Similarity 65.0%; Pred. No. 1e-90;
Matches 467; Conservative 0; Mismatches 243; Indels 9; Gaps 2;

QY 81 TCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 140
DB 74 TACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133

QY 141 GCCGAGCGATGCGCGGAGTGTGAACGCGTGAACAGTCTGACTGATTTTC 200
DB 134 GCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 193

QY 201 CCGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 260
DB 194 ACGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 253

QY 261 ATCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 320
DB 254 ATCATTTCTTAAGAAACCACTTTATTTATTCGCTTGGGAAAGTGAATGCTTCC 313

QY 321 ACGGCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 380
DB 314 ACGGCGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 373

QY 381 CTTTCTTGGTGGCGATTTTGGCTTATGATGATGATGATGATGATGATGATGATGAT 440
DB 374 ATTTTGGCAACGCTGATTTTGTGCGCTTGTGATTCGTTTGGTTAGGTA 433

QY 441 GTATCGTACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 500
DB 433 GATTTGTAATTTCAATTAATCTTGAATATGATGATGATGATGATGATGATGATGAT 493

QY 501 TTGTAATCCCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 557
DB 494 TGAATTTCCGATATATCTCTTTTACCTGCTGATGATGATGATGATGATGATGATGAT 553

QY 558 CTTTGAAGGCGTGTGCTGCTTCTTATTTTATTTATTTATTTATTTATTTATTTATTT 617
DB 554 TTTTGAAGGCGTGTGCTGCTTCTTATTTTATTTATTTATTTATTTATTTATTTATTT 613

QY 618 TAGGAGCGTATCCGAGCTGTTTATGATGATGATGATGATGATGATGATGATGAT 677


```

RESULT 7
US-09-252-991A-15227
; Sequence 15227 Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15227
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-15227

```

QY 389 GGGTGGCCGATTTTGTGTCGCCCTTTATGCGCATTCGGTTTGGGGGATG33ACATACGGT 448
Db 788 CAGCTATGAGCTTCAATGACCCCGCTGGTGTCCATTCGGCCTGGGCGCGCCGACATCGG 847
QY 449 AACTTATGATAGTGAACCTTGGGACGAGTAAACGATGTGCTTGCGCTTTGTATTC 508
Db 848 AACTCATCACTCGGAACCTGGGGGCAAGGTGACGATGTGCGCTTGCGCATGTGCTTC 907
QY 509 CCTAATGTGGCCCACTGCGCGGCATCTTCACAGCTTATGATGTGCGCTTAGAAGG 568
Db 908 CCCAAGGGGGCGCCGCTGCGCGGATCCTTGACAGCTGACATGTGCGCTTGAAAGG 967
QY 569 GTGGTCTGTCTTTATTTCTTAATTTGTTAATGTTAAGTAAACCTGTCGCTAGGACAGTA 628
Db 968 GTGGGCGCTGTGCTATCTCTGCGGTGTTACCGGCAAGCAGCGCGGCTCGGTGTC 102
QY 629 TCCGAGCTGTTTAACTGAGTACGATACATTCGCGCTTCCTGTGGAATACGTGCTGAG 688
Db 1028 TCCGCGCTGTGTGCTGTGCTAGCAACGAGATCTTCGCGCTTGTTGCAATTCGTCGCGCTG 108
QY 689 CCAGATGCTCAATGGG---TCTGTTTGGTGGCTTCAATTCATGGGGCAATCTCTGCG 745
Db 1088 CCGGATGCCAGCTGCGCTACCTGCGCTGGGGGTGGCTGACATGCGCCACAGTGTCTGCG 114
QY 745 TTACTATGTGATCATCGTATTTTGAATGATGTGTGTCTTACAAAGCGG 797
Db 1148 GTACGATGTGTGCGCGGACATTCGCTGATGTGCTGGGCGTAAACGCGCGG 1199

	Query Match	9.1%	Score 265.8	DB 4	Length 891
	Best Local Similarity	61.6%	Pred. No. 1.76-81		
	Matches	519	Conservative	0	Mismatches 252; Indels 72; Gaps 3
Qy	26	TATCTGCATTTTCCCATATTGACCCCGATTTTTCGATTCGGCCCTTCAGGGTGGCC	85		
Db	19	TACCTTAATTTTCGAAATCGATTCAGTATGTTTCATCGGGCCAGATTTTTCAC	78		
Qy	86	TGCTATGCTTGATGATTTTGGTGTTCTTTTTCATGATGATGGCCATTCGCCGA	145		
Db	79	TGCTACGGATGATGATTTGATTCGATTTGTGTTTCATTATGCGTAACCCGACA	138		
Qy	146	GCGGATCGCGCGGCGACGTGTGACGCGTGAACAATCTCTGACTGTTATCGCGGCG	205		
Db	139	GCGGCAAAACCTPAACAGTGTGTGGCGAAAAAGTGAAGTCGAAACATTTATGTAGGG	198		
Qy	206	TTTTTGTGTGATGATCGGTGGCGAATTGGTATGTGATCTTTCACAAATTTGATCTG	265		
Db	199	TTTGTGGGCTGTTTATGTGCGGGGTCTAGAGATATTTTATCTAATCACTTACCGGTA	258		
Qy	266	TTCTCTGCGAACCCCTTATTTATTCGAABGTGACGTGGCGCATGCTCTTCACGGG	325		
Db	259	TTTTGATTAATCCCTTACCTATTAAAGTCTGGAGATGCGGATGCTCTTCACGGT	318		

QY 326 GAGTATTGGGATGATCAGCCGATGTTGATGAGGCGTAATAACCAAGCAGCCTTC 385
DB 319 GCGTTAATGGGGATGATGCGCATATTTGTTTCTGCTGCACTAAACGTCCTTT 378
QY 386 TTTGGTGGCCGATTTTGTGCCCCCTTATGTCGCAATCGATTGGGAGTGGAGCATTC 445
DB 379 TTTAAGTCGCACTTTGTGCTCATTAATCCATTGTTAGGTTAAGACCATAT 438
QY 446 GGTAACTTATGATATGAGACTTTGGGAGCAGTAAC---GGATGGCCTTGAGCTTT 502
DB 439 GGTAACTTATTAATGAGAGCTTTGGGAGAGATCCTTAATACCCCGGAGATT 498
QY 503 GATTTCCC----- 510
DB 499 TTAATCCAGGCTCCCGTGCAGAAATCTACAGCTTGCAGCAAGTCCCACTACATA 558
QY 511 -----TAAATGGGCCCACTGGCGGCGCATCTTCACAGCTTATGAAATC 556
DB 559 TTACCTATCATTCAGCAATACGCGCATATGCTTGCACCAATGCACTCTATGAATG 618
QY 557 GCGTTAAGAGCGGTGCTGCTTCTTATTTATTTATGTTTATGTTAAACCTGCTCCG 616
DB 619 TTCCATGAAGCGGTGATATATTTATTTATTTGATATCTTTGGCGCAAAATCGTCT 678
QY 617 CTAGGCAAGCATCCGACCTGTTTATGCTGATAGCGTACATTCGCTTCTTGAGAA 676
DB 679 ATGGCAAGTATGAGGATTTCTTAATGTTATGAGTTCGCTTCTTAATTTATGAA 738
QY 677 TACGTCGTCGACGATGCTCAGTGGTCTGTTTGTGGCTTCATTTCAATGAGGCA 736
DB 729 TTTCTCCGCAACGATGATGCGCATATGATTTATTTAGCGTATCA---GCATGGGCA 795
QY 737 ATCCCTCTTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 796
DB 796 ATCTTCTCATTCATATGATATATATGAGCATATTTATGATGATGATGATGATG 855
QY 797 GGT 799
DB 856 GAT 858

RESULT 9
US-09-489-039A-2965
; Sequence 2965, Application US/09489039A
; Patent No. 66,0836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709,2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2965
; LENGTH: 879
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2965

Query Match 8.8%; Score 256.8; DB 4; Length 879;
Best Local Similarity 59.5%; Pred. No. 2,38-78;
Matches 515; Conservative 0; Mismatches 282; Indels 69; Gaps 2;

QY 14 ATGCTTCAGGCTATCTGCAATTTCCATATTTGACCCCGTATTTTTCGATCGGCT 73
DB 4 ATGAATAGTGGTACTGCTGATTTTCGAGATTCATCGGCAATTTTCTCTAGGCGCG 63
QY 74 CTAGCGGTGCGGTGATGCTGATGATGATGATGATGATGATGATGATGATGATG 133
DB 64 GTCTCCCTCCACTGATGATGATGATGATGATGATGATGATGATGATGATG 123

QY 134 GCGAATGCGCGAGGATGATGCGGCGGAGTGTGAGCGGCTGAGCAAGTCTGACTG 193
DB 124 GCGACCGGTGCGGAGATTCGCGGAGCGGCTGAGCAAGTGTGAGCAAGTCTGACTG 183
QY 194 TTAATGCGCGCTTTTATGATGATGATGATGATGATGATGATGATGATGATG 253
DB 184 CTGACCGCGGCTTCCTCGCGCTTCTGCGGCGGCTGATGATGATGATGATGATG 243
QY 254 AATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 313
DB 244 AACCTGCGGATTTCTTCCGATTCGCTGATCTGCTGCTGCTGCTGCTGCTGCTG 303
QY 314 TCTTCACAGCGCGCTTATGATGATGATGATGATGATGATGATGATGATGATG 373
DB 304 TCGTTCAAGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 363
QY 374 CAAGCAAGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 433
DB 364 AAAGCAAGCTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 423
QY 434 ATGGAAGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 481
DB 424 GCGGCGGTGCGGCACTTATCAAGCGGAGCTGAGGCGCGTACAGCGGCTTC 483
QY 482 -----ACGAT 487
DB 484 CACTACAGATATTTCCCGGCTCCCGCGGAGAGCTGAGCGCTGCGGAGCAT 543
QY 488 GTGCTTGGCTTTTATTCCTAATGATGATGATGATGATGATGATGATGATGATG 547
DB 544 CCGAATGCGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 603
QY 548 TATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 607
DB 604 TACAGCTGCGGCTGAGGAGGATGATGATGATGATGATGATGATGATGATGATG 663
QY 608 CCGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 667
DB 664 CCGGCGCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
QY 668 CTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 724
DB 724 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 783
QY 725 TCAATGCGGCAATCTCTCTTACTATGATGATGATGATGATGATGATGATGATG 784
DB 784 AGCATGCGGCAATCTCTCTTACTATGATGATGATGATGATGATGATGATGATG 843
QY 785 TCTTCAAGCGGCTTGTATCAAGA 810
DB 844 GCTTACCGCAGCGTCCGACAGCA 869

RESULT 10
US-09-328-352-3810
; Sequence 3810, Application US/09328352
; Patent No. 66,2858
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-039A
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3810
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3810

Query Match 7.2%; Score 208.6; DB 4; Length 834;

Best Local Similarity 58.0%; Pred. No. 1.4e-61;
Matches 452; Conservative 0; Mismatches 309; Indels 18; Gaps 4;

Qy	29	CTGCAAGTTCCCAATATGACCCCGATATGTTTTCGATGCGCCCTCAGCGGCGCTG	88
Db	19	CTGACCCATCCCAATATGATCCGGTCGCAATCACTAGACACCTCTTCAAGTCAATGG	78
Qy	89	TATGCTTATCATATTGGTGGTTTCCTTTTGTATGTGTGGCAATCCGCGAGG	14
Db	79	TATGACCTCATGTAATTATGGCAATTTTATGTGCTGGGGCTGTCTCTACCGTGC	13
Qy	149	GATCGCGCGGACAGGTGTGACCGCGTAGCAAGTCTGACTTGTATTTCGCGCGCTT	20
Db	139	AAACA---GCGTATGGCTGGACATGGAGCATGGTTCCGATCTGGTGTTTTACGGCGC	19
Qy	209	TTAGGTATGATATCGGTGGCCGAGTTGGTTATGTGATCTTCTCAATTTTATCTGTTC	26
Db	196	TATAGTATTGACTCGGGAGCGATTCGGCTATGTGCTTTCTATGAGTTGATAAATTC	25
Qy	269	CTTGTGCACCCCTCTTTATTTATTCAAAGTGTGACCTGGCGGACGATGCTTCCAGCGGCG	32
Db	256	CTTGAAAACCCATCTGTGTTATTCGAAAGTCTGGACAGGTGTATGAGCTTCACGCGGCG	31
Qy	329	TTATTTGGGTGTATCACCGCAGTTCGTGGTATCGCGCTAAAAACAACGACCTTCCTT	38
Db	316	TTCTTATGCGGTCAATGATTCGCAATGCTATCTGTGTAAATAATCTAAATAACATGGTTC	37
Qy	389	GGTGTGCGCGAATTTGTGTGCCCCCTTATGTCGCAATTCGGTTTGGGATGGGACGTATCGGT	44
Db	376	CAGACATCTGACTTTGTGTGCCCTTGTGTACCGACAGGCTTATATCTTGGACGTATGGGT	43
Qy	449	AACCTTATGAAATAGTGAATTTGGGAGCG-----AGTACGAGATGTGCTTGGGCT	49
Db	436	AACCTTATTTGGCGAGAGTTATATGGTCTGTGGGTACAGATCCGAATTCATCATTCGGA	49
Qy	500	TTTGTATTCCCTA---ATGTGTGCCACAGCGCGGCGCAATCCTTCAGAGCTTTATGAATTC	55
Db	496	ATGATTTTCCCAACCGATCATATKATCTGGTGTGTACACCTTCTCAATCTACCAAGG	55
Qy	557	GCCTTAAGAGCGGTGTTCTGTCTTTATCTTAATATGTTTATTTGTAACCTGTGCG	61
Db	556	CTTTTCGGAAGACATATTAATCTTTATTTATTTTGTGTGTGTACCTAAACACGCGCCA	61
Qy	617	CTTAGCAGCGATATCCGAGCTTTTATAGTGTATACGCTAATTCGCGTCCCTGTGGA	67
Db	616	CGTATGGCTGATCTGCGCCATTTTTTAATAGGTTATATGTGTGACGCGTTTGTGATGAA	67
Qy	677	TACGTGCGTAGCCAGATGCTCACTTGGGT---CTGTTGTGGCTTCATTCATATGGG	73
Db	676	TTCTTCGCTACGCGGATTCGACCAAGGTTTCATCTGTTTGGCTGAGTACCAAGGA	73
Qy	734	CAATACCTCTTAACCTATATGTATGATCATCGTATTTTGAATAGGTTTGTGTTACA	792
Db	736	CAAAATCTCAACCTTCCAAATGCTACTGATAGGCGTTTGAATGATGTGTATGCTTACCA	794

```

RESULT 11
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632835
; GENERAL INFORMATION:
;   APPLICANT: SHIGENOBU, SHUJI
;   APPLICANT: MATENABE, HIDEMI
;   APPLICANT: HATTORI, MASHIYUKI
;   APPLICANT: SAKAKI, YOSHITSUKU
;   TITLE OF INVENTION: GENEOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;   FILE REFERENCE: 081356/0159
;   CURRENT APPLICATION NUMBER: US/09/790,988
;   CURRENT FILING DATE: 2001-02-23
;   PRIOR APPLICATION NUMBER: JP2000-107160
;   PRIOR FILING DATE: 2000-04-07
;   NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp
US-03-790-988-1

```

Score 152.2; DB 4; Length 640681,
Pred. NO. 1 3e-39;

Matches 432; Conservative 0; Mismatches 323; Indels 66; Gaps 2;

QY	35	TTTTCCCAATATACCCCGGATATGTTTCGATCGGCGCCTTAGCGGATGCGGTATAGC	94
Db	479304	TTTCCATAATTAATCCATAATTTTACGATTTGGCTGATCTGCTGCGGTATAGT	479363
QY	95	TTGATGATTTGGGTGGCTTCCCTTTTGGTATGTTGGTTGGCCATCCGCGGAGATCCG	154
Db	479364	TTTATGATGTTATTAAGTTTATTTATTTGTAATGATGTGAGAAAATACTGCATATTAA	479422
QY	155	GCGGGCAGTGTGTGACCGGTAGCAGATCTCTGACTTGTATTTCGCGGCTTTTAAAGT	214
Db	479424	AATTAATAAATATGATGACAAAAAATAAATAGAAAATATTATTATTTCTATTTTATGGA	479483
QY	215	GTATGATACGGTGGCCGAGTTGGTATATGATATCTCTACATTTTGAATCTGTCCCTGCT	274
Db	479484	TCAAGCATGGAGAAAGAAATAGATATATTATTTTTAATTTTCAATATTATTTCTGA	479543
QY	275	GACCCCTTATTATTATTTCAAGTGTGACGTGCGGCAATGCTTCCACGCGGCTTATG	334
Db	479544	AATATGTTATCTGATTTATTATATATGAGAGAGGCAATGATTCATATGAGAGATTATA	479603
QY	335	GGTGTATACCGGCAATGTTCTGTATATGCGGTAAACCAAGCACCTTTTGGGTG	394
Db	479604	GGTGCATATGCTCATATGCTGATTTTCTTTAATATATAAAAAATAATTTGAAATA	479663
QY	395	GCCGATTTTGTGCCCCCTTTAGTGCATTCGCTTGGGATGGACGTAATCGTACTTT	454
Db	479664	TCTGATTTATCATCTCCACTAATATACCTTTTGGTTAGCGCGGTAAATAGCAAACTTT	479723
QY	455	ATGAATAGGAATCTTGGGAGAGAGT---AAGGATGAGCTTGGGCTTTGATTCCT	511
Db	479724	ATTAAATAGTAACTATGGGTGTGTATACCAATTTTTCATATGCAATGATTTTCCC	479783
QY	512	AATGTG-----	517
Db	479784	AATTCCAAAATCAAGATTTAAAGAAATATAAAAAATATCCGGAATTACACTATTATTA	479843
QY	518	-----GGCCACTGCGCGGCATCTTCACAGCTTATGSAATTGCGCTTAGAAGGC	568
Db	479844	GATCAATACGAGACATTCGCGGTATCTTACCAATATATCGAAATTTTTTTGAAGGG	479903
QY	569	GTGGTCTGTCTTATTCTTTATTATGTTTATATGTTTATATGTTAACTCGTCCGTAGGCA	628
Db	479904	ATTCTCTGTCTTTTATATTAATTAATTTTTCGAAAAAGATAGCCAAAGGATGATTT	479963
QY	629	TCCGAGCTTTTATAGCTGATAGAGGTACATTCGCGCTCCTGTGAGAAATGCTCCGAG	688
Db	479964	AGCGGTTTATTTTATTTTATTTTATGATATATTAGAAATATTATAGAAATTTTTAGGA	480023
QY	689	CCAGATGCTCACTGGTCTGTTGGTGGCTTCAATTCATATGGGCAAAATCTCTCTTA	748
Db	480024	CCAGATCCCAAAATAGACTATTAATAAAATATTAATTTCTATAGGACAAATATTATAC	480083
QY	749	CCATATGATATCATCGATTTTATGATATGTTGGTTGGCTTA	789
Db	480084	CCATATATTTGCTCGATTAATTTATATGATTAATCTTA	480124

RESULT 12
US-09-540-236-663
; Sequence 663, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540.236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO: 663
LENGTH: 891
TYPE: DNA
ORGANISM: M.catarrhalis
US-09-540-236-663

Query Match 4.9%; Score 143.8; DB 4; Length 891;
Best Local Similarity 59.4%; Pred. No. 5.4e-39;
Matches 263; Conservative 0; Mismatches 177; Indels 3; Gaps 1;

38 CCCAATATTGACCCCGATTTGTTTCATGCGCCCTTCAGCCGCTGCGCTGATGAGCTTG 97
28 CCACGATTAACCTGTTGCTTTAGCTTGGGCATGATAGCTGCTGCTGATGCTTG 87
98 ATGATTTGGTGGCTTCTTTTGTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 157
88 ATGATTTGCTTCTTTTGTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 144
158 GCGAGTGTGAGCGCGTGAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 217
145 CGAACAGATTTTACCCCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 204
218 GTGATCGGTGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 277
205 ATTTAGTGGCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 264
278 CCTCTTATTTATCAAGTGTGAGCTGCGCGATGCTGCTGCTGCTGCTGCTGCTG 337
265 CCTTATATCTTTTTCAGTATGGAAGGCGCATGCTGCTGCTGCTGCTGCTGCTG 324
338 GTGATCAACCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 397
325 GTGTTGCTGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 384
398 GATTTTGTGCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 457
385 GATTTATCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 444
458 AATAGTGAATTTGGGCGAGCT 480
445 AATGCGAGCTTTGGGCGAGAT 467

RESULT 13
US-09-596-002-35
Sequence 35, Application US/09596002
Patent No. 6632636
GENERAL INFORMATION:
APPLICANT: Lagace, Robert, E.
APPLICANT: Paterson, Chandra
APPLICANT: Berg, Kim, L.
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
FILE REFERENCE: PM-0008-4 US
CURRENT APPLICATION NUMBER: US/09/596.002
CURRENT FILING DATE: 2000-06-16
PRIOR APPLICATION NUMBER: 60/140,121
PRIOR FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PERL Program
SEQ ID NO 35
LENGTH: 96109
TYPE: DNA
ORGANISM: M. catarrhalis
FEATURES:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte template ID No. 6632636 35
PUBLICATION INFORMATION:

US-09-596-002-35

Query Match 4.9%; Score 142.2; DB 4; Length 96109;
Best Local Similarity 59.1%; Pred. No. 8.3e-37;
Matches 262; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

38 CCCAATATTGACCCCGATTTGTTTCATGCGCCCTTCAGCCGCTGCGCTGATGAGCTTG 97
Db 51130 CCACGATTAACCTGTTGCTTTAGCTTGGGCATGATAGCTGCTGCTGCTGCTG 51189
Qy 98 ATGATTTGGTGGCTTCTTTTGTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 157
Db 51190 ATGATTTGCTTCTTTTGTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 51246
Qy 158 GCGAGTGTGAGCGCGTGAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 217
Db 51247 CGAACAGATTTTACCCCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 51306
Qy 218 GTGATCGGTGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 277
Db 51307 ATTTAGTGGCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 51366
Qy 278 CCTCTTATTTATCAAGTGTGAGCTGCGCGATGCTGCTGCTGCTGCTGCTGCTG 337
Db 51367 CCTTATATCTTTTTCAGTATGGAAGGCGCATGCTGCTGCTGCTGCTGCTGCTG 51426
Qy 338 GTGATCAACCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 397
Db 51427 GTGTTGCTGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 51486
Qy 398 GATTTTGTGCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 457
Db 51487 GATTTATCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 51546
Qy 458 AATAGTGAATTTGGGCGAGCT 480
Db 51547 AATGCGAGCTTTGGGCGAGAT 51569

RESULT 14
US-09-252-991A-15244/C
Sequence 15244, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 35142
SEQ ID NO 15244
LENGTH: 435
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15244

Query Match 4.0%; Score 115; DB 4; Length 435;
Best Local Similarity 59.2%; Pred. No. 3.3e-29;
Matches 196; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

29 CTGAGTTTCCCATATTGACCCCGATTTGTTTCATGCGCCCTTCAGCCGCTGCGCTG 88
Db 373 CTGAGTTTCCCATATTGACCCCGATTTGTTTCATGCGCCCTTCAGCCGCTGCGCTG 314
Qy 89 TATGCTTGAATTTGTTGGGCTTCTTTTGTATGCTATGCTGCTGCTGCTGCTGCTG 148
Db 313 TACGCGTGAATTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 254

QY 149 GATCGCGCGGAGTGTGGACGCGTGAAGTCTCTGACTTGTATTGCGCGGCTTT 208
 DB 253 AGCGCTTGAACCCGACCTGACCAAGAGCGCTTCCGACCTGCTCTTGGGTGCGC 194
 QY 209 TTAGGTGTAGTATGCGGCGGAGTGTGTATGTATCTTCAACAATTTGATCTGTC 268
 DB 193 TCGGCGGTATCTCTGGGCGCGCTGCGCTACGCTGCTTCTTCAACCTGGAAGGTC 134
 QY 269 CTGTGTGACCCCTCTTATTATTATGAAAGTGTGAATGCGCGCAATGCTTCCAGCGCGC 328
 DB 133 ATCGCCAACCCGACGCTGATCTTGAAGGTGTGAAGGCGGCAATGCTTCCATGCGCGC 74
 QY 329 TTAGGTGTAGTATGCGGCGGAGTGTGTATGTATCTTCAACAATTTGATCTGTC 359
 DB 73 CTGCTGGCGGTATCTCTGGCGCTCTGCTGCT 43

RESULT 15
 US-08-935-433-1/C

; Sequence 1, Application US/08935433

; Patent No. 6319688

; GENERAL INFORMATION:

; APPLICANT: FELLD, JOHN

; TITLE OF INVENTION: A HUMAN SODIUM DEPENDENT PHOSPHATE

; TITLE OF INVENTION: TRANSPORTER (IPT-1)

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER & PRESTIA

; STREET: P. O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/935,433

; FILING DATE: 23-SEP-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/044,974

; FILING DATE: 28-APR-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: PRESTIA, PAUL F

; REGISTRATION NUMBER: 23,031

; REFERENCE/DOCKET NUMBER: GH-70006

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-407-0700

; TELEFAX: 610-407-0701

; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2288 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; US-08-935-433-1

Query Match 2.6%; Score 75.4; DB 4; Length 2288;
 Best Local Similarity 56.0%; Pred. No. 7.8e-15;
 Matches 164; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 2476 CATTTCTAGAGGTAGAGATCAGACCGCAGAGAGATTGAAAAATCGTAGTGTGGC 2535
 DB 771 CACCTCCACGGGCAAGAGACCAACAGCAGCAGTGAAGATCATGACAGTGGC 712
 QY 2536 ACTGGCGAAGACGCGCAAACTCTTCTTACAGCGCATATGCGCAAGCGTGAAGAT 2595
 DB 711 TCTGTAAAGCTTTCTTGAATCACTCCGATCTTCCACCTGCAATGAGCGCAATAGT 652

QY 2596 ATTGTCACAGTAGTACCAATATTTGGACCCATCAACCATAGGAATCGCGTTTCAACG- 2654
 DB 651 GTTGATGATGACGTTTCAATGTTGCCCCCATGATTAATGGGAGTGGACCGCAACAGT 592
 QY 2555 --GTAAACCCACCGGCAACGAGACCAACATAATAGAGTCAACCGCTTGAAGATTGAAT 2712
 DB 591 GAGCAATGAAAGAGACACCATGCTGACACGATGACGTTGAGGTGCTGAGCTGTGAC 532
 QY 2713 CAGTCCGTTGCCACTTAAACCAATCATCATCTTGCATTTGGGTGGAAACCA 2765
 DB 531 CAAGACGTCACGACACCCCGATCAACGCCCCCAACAAAGGTTGACATTA 479

Search completed: July 31, 2004, 17:02:39
 Job time: 160 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 07:39:00 ; Search time 7452 Seconds

(without alignments)
16919.604 Million cell updates/sec

Title: US-09-700-712A-1

Perfect score: 2909
Sequence: 1 gagaaggttctgtatgcctc.....cagcgaatcgcgctgcag 2909

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

GenEmbl: 1: gb_ba: 2: gb_hbg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_mu: 20: em_om: 21: em_ov: 22: em_ph: 23: em_pat: 24: em_pi: 25: em_ro: 26: em_sts: 27: em_un: 28: em_vl: 29: em_hbg_hum: 30: em_hbg_inv: 31: em_hbg_oth: 32: em_hbg_mus: 33: em_hbg_dyn: 34: em_hbg_rod: 35: em_hbg_mam: 36: em_hbg_vit: 37: em_hbg_hum: 38: em_hbg_mus: 39: em_hbg_oth: 40: em_hbg_mus: 41: em_hbg_oth: 42: em_hbg_oth: 43: em_hbg_oth: 44: em_hbg_oth: 45: em_hbg_oth:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2909	100.0	2909	1	VCH6514
2	2909	100.0	2909	6	AX009481
3	2870.6	98.7	9931	1	AE004153
4	1390.8	47.8	1410	1	VCH7135
5	1222	42.0	1222	6	AX009483
6	1216.8	41.8	1236	1	AY143429
7	1187	40.8	1591	1	VCA10968
8	1052.4	36.2	298900	1	AP005074
9	1035	35.6	301442	1	AE016798
10	1028.6	35.4	249150	1	AP005332
11	841.4	28.9	852	6	AX020429
12	838	28.8	838	6	AX009482
13	814.8	28.0	10719	1	AE006043
14	775.2	26.6	1893	1	AF064791
15	744.2	25.6	11545	1	U12772
16	744.2	25.6	110000	6	AR274543_09
17	433.8	14.9	300521	1	AE017153
18	410.8	14.1	873	6	AR377276
19	352.2	12.1	304250	1	EX294452
20	329.4	11.3	10657	1	AE015577
21	314.8	10.8	304558	1	AE017151
22	292.2	10.0	807	6	AR139598
23	280.8	9.7	10474	1	AE004472
24	280.6	9.6	10029	1	AE013918
25	280.6	9.6	208050	1	AE141414
26	278.2	9.6	349960	1	EX571861
27	278.2	9.4	10242	1	AE005511
28	273	9.4	26658	1	AP002563
29	265.8	9.1	891	6	AR377359
30	264.2	9.1	305325	1	AE016765
31	261.4	9.0	310613	1	AE016875
32	260.8	8.9	10405	1	AE000366
33	260.2	8.9	12943	1	AE015288
34	260.2	8.9	292906	1	AE016968
35	259.6	8.9	71128	1	ECU29581
36	258	8.9	1473	1	STYLGIX
37	258	8.9	21583	1	AE008838
38	256.8	8.8	879	6	AR386236
39	256.4	8.8	230050	1	AL627277
40	256.4	8.8	301311	1	AE016843
41	256.4	8.5	876	6	AX188961
42	248.2	8.5	1310	1	ECOLGT
43	248.2	8.3	301086	1	AE016964
44	240.8	8.0	303438	1	AE017212
45	232.6	8.0	303438	1	AE017212

ALIGNMENTS

RESULT 1
VCH6514
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

VCH6514
Vibrio cholerae lgt and thya genes.
AY006514
AY006514.1 GI:3201563
lgt gene, prokaryotic diacylglycerol transferase; thya;
thymidylate synthetase.
Vibrio cholerae
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.

Carlén, N. A., Nilsson, A., Todorovic, M., Holmgren, J., and Lebens, M.
Characterization of the lgt/thya locus from Vibrio cholerae

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2909)
 AUTHORS Carlin N.I.A.
 TITLE Direct Submission
 JOURNAL Submitted (29-Mar-1998) Carlin N.I.A., Department of Molecular
 Biology, SBL Vaccin AB, PMB, UV 30, Stockholm, S-105 21, SWEDEN

FEATURES
 source
 1. 2909
 /organism="Vibrio cholerae"
 /mol_type="genomic DNA"
 /strain="JS1569"
 /db_xref="taxon:666"

gene
 14..829
 /gene="igt"

CDS
 14..829
 /gene="igt"
 /codon_start=1
 /transl_table=1
 /product="Proteoglycan diacylglycerol transferase"
 /protein_id="CA07072.1"
 /db_xref="GI:3201564"
 /db_xref="GOA:O87077"

gene
 669..775
 /gene="thyA"
 /transl_table=1
 /product="Methylglutamate synthetase"
 /protein_id="CA07073.1"
 /db_xref="GI:3201565"
 /db_xref="GOA:O66108"

-35_signal
 669..775
 /gene="thyA"
 -10_signal
 691..696
 /gene="thyA"
 gene
 join(828..832,839..1690,1725..1765)
 /gene="thyA"
 828..832
 /gene="thyA"
 839..1690
 /gene="thyA"
 /codon_start=1
 /evidence="experimental"
 /product="thymidylate synthetase"
 /protein_id="CA07073.1"
 /db_xref="GI:3201565"
 /db_xref="GOA:O66108"
 /db_xref="SWISS-PROT:O66108"

CDS
 839..1690
 /gene="thyA"
 /codon_start=1
 /evidence="experimental"
 /product="thymidylate synthetase"
 /protein_id="CA07073.1"
 /db_xref="GI:3201565"
 /db_xref="GOA:O66108"
 /db_xref="SWISS-PROT:O66108"

variation
 1082..1084
 /gene="thyA"
 /note="in trimethoprim resistant mutant JS1569 4.4, scop
 codon"

stem_loop
 1735..1765
 /gene="thyA"

ORIGIN
 Query Match 100.0%; Score 2909; DB 1; Length 2909;
 Best local similarity 100.0%; Pred. No. 0;
 Matches 2909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGAGGTTTGTATGCTCAGGTTATCTGAGTTCCCAATATTGACCCCGATTGTT 60
 DB 1 GAGAGGTTTGTATGCTCAGGTTATCTGAGTTCCCAATATTGACCCCGATTGTT 60
 QY 112 GATGCGGCTTACCGGCTGCGTGTATGCTGATGATTTGGGGGTTTCTTT 120

DB 61 TTGATGCGGCTTACCGGCTGCGTGTATGCTGATGATTTGGGGGTTTCTTT 120
 QY 121 TTGATGCGGCTTACCGGCTGCGTGTATGCTGATGATTTGGGGGTTTCTTT 180
 DB 121 TTGATGCGGCTTACCGGCTGCGTGTATGCTGATGATTTGGGGGTTTCTTT 180
 QY 181 AGTCTGACTGTTTATGCGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 240
 DB 181 AGTCTGACTGTTTATGCGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 240
 QY 241 TGTGATCTTCTAATTTTATGCTGATGATTTGGGGGTTTCTTT 300
 DB 241 TGTGATCTTCTAATTTTATGCTGATGATTTGGGGGTTTCTTT 300
 QY 301 GACTGCGGCTTCTTACCGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 360
 DB 301 GACTGCGGCTTCTTACCGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 360
 QY 361 TGGCGCTTAAACCAACGACCTTTCTTGGTGTGCGGATTTGGGGGTTTCTTT 420
 DB 361 TGGCGCTTAAACCAACGACCTTTCTTGGTGTGCGGATTTGGGGGTTTCTTT 420
 QY 421 ATTGCGTTGGGGATGCGATGCTGATGATTTGGGGGTTTCTTT 480
 DB 421 ATTGCGTTGGGGATGCGATGCTGATGATTTGGGGGTTTCTTT 480
 QY 481 AACGATGCGCTTGGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 540
 DB 481 AACGATGCGCTTGGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 540
 QY 541 AACGATGCGCTTGGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 600
 DB 541 AACGATGCGCTTGGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 600
 QY 601 TGTGATGCGCTTGGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 660
 DB 601 TGTGATGCGCTTGGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 660
 QY 661 CCGCTTCTTGGGATGCGATGCTGATGATTTGGGGGTTTCTTT 720
 DB 661 CCGCTTCTTGGGATGCGATGCTGATGATTTGGGGGTTTCTTT 720
 QY 721 CATTTCAATGCGGCAATCTCTCTTACCTTATGCTGATGATTTGGGGGTTTCTTT 780
 DB 721 CATTTCAATGCGGCAATCTCTCTTACCTTATGCTGATGATTTGGGGGTTTCTTT 780
 QY 781 TTGCTCTTCAAGCGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 840
 DB 781 TTGCTCTTCAAGCGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 840
 QY 841 GAAACAGTATTAGATCTTTGTCAGCGGATGCTGATGATTTGGGGGTTTCTTT 900
 DB 841 GAAACAGTATTAGATCTTTGTCAGCGGATGCTGATGATTTGGGGGTTTCTTT 900
 QY 901 ACGAAGCGGCAAGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 960
 DB 901 ACGAAGCGGCAAGGCTTTTATGCTGATGATTTGGGGGTTTCTTT 960
 QY 961 CAATCAATTTCTCTAGTACTTACCAACGAGATTTTGGAAAGCTGCGTACCGAGTT 1020
 DB 961 CAATCAATTTCTCTAGTACTTACCAACGAGATTTTGGAAAGCTGCGTACCGAGTT 1020
 QY 1021 GCTGCGCTTATGCTGCTTACCAATGCGGCGGATTTTGGCAATATGATGACCAAAAC 1080
 DB 1021 GCTGCGCTTATGCTGCTTACCAATGCGGCGGATTTTGGCAATATGATGACCAAAAC 1080
 QY 1081 CTGGATGCTTATGCTGCTTACCAATGCGGCGGATTTTGGCAATATGATGACCAAAAC 1140
 DB 1081 CTGGATGCTTATGCTGCTTACCAATGCGGCGGATTTTGGCAATATGATGACCAAAAC 1140
 QY 1141 GATGATGCTTATGCTGCTTACCAATGCGGCGGATTTTGGCAATATGATGACCAAAAC 1200
 DB 1141 GATGATGCTTATGCTGCTTACCAATGCGGCGGATTTTGGCAATATGATGACCAAAAC 1200

QY 1201 TCATATTGACAGTTGAAAAAGATTGTGATGTTTGAAGCGGTGGCTTGATGACCGAGG 1260
 Db 1201 TCATATTGACAGTTGAAAAAGATTGTGATGTTTGAAGCGGTGGCTTGATGACCGAGG 1260
 QY 1261 TGAATATCTTAACCTTACATCCGGGTGAAATTCACATGGGGTGTTCGGCCCTTGAT 1320
 Db 1261 TGAATATCTTAACCTTACATCCGGGTGAAATTCACATGGGGTGTTCGGCCCTTGAT 1320
 QY 1321 GTACAGCCATCATTTTTCATGCTGGGGATACCTTGATCTTCAACAGTACAGGTTTC 1380
 Db 1321 GTACAGCCATCATTTTTCATGCTGGGGATACCTTGATCTTCAACAGTACAGGTTTC 1380
 QY 1381 ATGTGATGTCCTTGGGGTGAATTTCAACATGTCAGGTTTATGTCCTTGCGGT 1440
 Db 1381 ATGTGATGTCCTTGGGGTGAATTTCAACATGTCAGGTTTATGTCCTTGCGGT 1440
 QY 1441 GATGCAACAGATCACAGGGAAAAAGCCGGGCTTGCGTATCAACAGATCGTCAATGCCA 1500
 Db 1441 GATGCAACAGATCACAGGGAAAAAGCCGGGCTTGCGTATCAACAGATCGTCAATGCCA 1500
 QY 1501 CATTTACCAAGATCACTCGAATTTGATGCGCATGTCAGCTTAAACCTGACCATTCCTC 1560
 Db 1501 CATTTACCAAGATCACTCGAATTTGATGCGCATGTCAGCTTAAACCTGACCATTCCTC 1560
 QY 1561 AGCGCTCAAGTTCATATCATCAAAAGATTAAACACTGACAGATTGGAATCTTGAGT 1620
 Db 1561 AGCGCTCAAGTTCATATCATCAAAAGATTAAACACTGACAGATTGGAATCTTGAGT 1620
 QY 1621 CACTTGGATGATTTTGAAGTCAACGGATATCACTTCAAGATCTTATTCATACCTGTT 1680
 Db 1621 CACTTGGATGATTTTGAAGTCAACGGATATCACTTCAAGATCTTATTCATACCTGTT 1680
 QY 1681 TTCACTATCCCGTATTCAGGCGGTATGCTGATGGGTTTATATATAAAAGCTCC 1740
 Db 1681 TTCACTATCCCGTATTCAGGCGGTATGCTGATGGGTTTATATATAAAAGCTCC 1740
 QY 1741 CGAAGTCGGAAGCTTTTATATACAGATGATGCTTAAACGCTTAAAGCGTTAGGCAAG 1800
 Db 1741 CGAAGTCGGAAGCTTTTATATACAGATGATGCTTAAACGCTTAAAGCGTTAGGCAAG 1800
 QY 1801 AATGCTGCGGGATGACGAAACACACCAATTAAGTACCTACACCAATTTGCT 1860
 Db 1801 AATGCTGCGGGATGACGAAACACACCAATTAAGTACCTACACCAATTTGCT 1860
 QY 1861 CTTACAGCCCAAGTGAATGATGCTCAGACCTTTAATAGGCACTTCGCTAAGAAAG 1920
 Db 1861 CTTACAGCCCAAGTGAATGATGCTCAGACCTTTAATAGGCACTTCGCTAAGAAAG 1920
 QY 1921 AATACCGTAAATCAAGACCGTGAACCTCAAGTTAAAGCTTAAAGTGACCAAGCAATTTG 1980
 Db 1921 AATACCGTAAATCAAGACCGTGAACCTCAAGTTAAAGCTTAAAGTGACCAAGCAATTTG 1980
 QY 1981 CAGAGCAACACGCGAACTCAACAGACAGCGGTTGCGGAGAGAGAGAGAGAGTAAACA 2040
 Db 1981 CAGAGCAACACGCGAACTCAACAGACAGCGGTTGCGGAGAGAGAGAGAGAGTAAACA 2040
 QY 2041 AGTCCCAATGTTGCGACCTTAAGGTAAATGGGTAGATTTCAGCAGCTTTACAGCGCCAGA 2100
 Db 2041 AGTCCCAATGTTGCGACCTTAAGGTAAATGGGTAGATTTCAGCAGCTTTACAGCGCCAGA 2100
 QY 2101 GCCCAAGAGAGAACCATTAAGGCTGTTGTCGATGAGTGAAGTTAATCAATCAATCCGTAC 2160
 Db 2101 GCCCAAGAGAGAACCATTAAGGCTGTTGTCGATGAGTGAAGTTAATCAATCAATCCGTAC 2160
 QY 2161 CACTGACTGAAGCAATACCGTGTAGTGGCTCGGCAATCGCATTTTGTAGATTTTC 2220
 Db 2161 CACTGACTGAAGCAATACCGTGTAGTGGCTCGGCAATCGCATTTTGTAGATTTTC 2220
 QY 2221 AGGTGGGGGCAACCATTAATCTTCACTGATGTTTGGCCATCACCGTAAATGGCAGCA 2280
 Db 2221 AGGTGGGGGCAACCATTAATCTTCACTGATGTTTGGCCATCACCGTAAATGGCAGCA 2280

QY 2281 AATGTCCTCAATACCCATATACGATAGTGGCAACACCCAGAAATATTACCAATACCGA 2340
 Db 2281 AATGTCCTCAATACCCATATACGATAGTGGCAACACCCAGAAATATTACCAATACCGA 2340
 QY 2341 AAGCTGGGTTTCAAGCCCTGTGATGACAGGTTTGGTATCGGTTGATTAATCAAAACC 2400
 Db 2341 AAGCTGGGTTTCAAGCCCTGTGATGACAGGTTTGGTATCGGTTGATTAATCAAAACC 2400
 QY 2401 TTTCATGCTCAATGCGCAGTGGCAAGAGGCGAAACGAGCCAGTGTGAGACTTTCTC 2460
 Db 2401 TTTCATGCTCAATGCGCAGTGGCAAGAGGCGAAACGAGCCAGTGTGAGACTTTCTC 2460
 QY 2461 TAAATGCAACATCAATTTCTAGAGTGAAGAAATACAGACCCCGGAGAAATGAAAAA 2520
 Db 2461 TAAATGCAACATCAATTTCTAGAGTGAAGAAATACAGACCCCGGAGAAATGAAAAA 2520
 QY 2521 ATCGTGAATGTCACCTGGCGAAAGCAACGCGGAACTCTTTTACAGCCATATGCGC 2580
 Db 2521 ATCGTGAATGTCACCTGGCGAAAGCAACGCGGAACTCTTTTACAGCCATATGCGC 2580
 QY 2581 AAGCTGACGAGATATTGCTCAAGATGTAACCAATATTGGCAACCATACATAGGAAT 2640
 Db 2581 AAGCTGACGAGATATTGCTCAAGATGTAACCAATATTGGCAACCATACATAGGAAT 2640
 QY 2641 CGCGGTTTCAACCGGTAACCAACCGGCAACGAGCAACCAATATAGAAAGTCAACGCTGCT 2700
 Db 2641 CGCGGTTTCAACCGGTAACCAACCGGCAACGAGCAACCAATATAGAAAGTCAACGCTGCT 2700
 QY 2701 TGAGATTTGAATCAAGTCGCGGTGCGCACTAAACCAATCATCATATCTGCAATTGGGTGGA 2760
 Db 2701 TGAGATTTGAATCAAGTCGCGGTGCGCACTAAACCAATCATCATATCTGCAATTGGGTGGA 2760
 QY 2761 AGCAATTTCAATGAATCTTGGCTGATGCGCGGTGGCCATTAAACCGCTGCCGAC 2820
 Db 2761 AGCAATTTCAATGAATCTTGGCTGATGCGCGGTGGCCATTAAACCGCTGCCGAC 2820
 QY 2821 CATGCGACTGCAAGAAATGATTAATACAGATGAAGAAAGCAAGTTGCCCAAGCTAGGCC 2880
 Db 2821 CATGCGACTGCAAGAAATGATTAATACAGATGAAGAAAGCAAGTTGCCCAAGCTAGGCC 2880
 QY 2881 TTTGCTGCTCAGCGAATTCGCGCTGCGAG 2909
 Db 2881 TTTGCTGCTCAGCGAATTCGCGCTGCGAG 2909

RESULT 2
 AX009481
 LOCUS 2909 bp DNA linear PAT 06-SEP-2000
 DEFINITION Sequence 1 from Patent WO9961634.
 ACCESSION AX009481
 VERSION AX009481.1 GI:9996766
 KEYWORDS
 SOURCE
 ORGANISM
 Vibrio cholerae
 Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 Vibrionaceae; Vibrio.

REFERENCE
 1
 Carlin, N. and Lebens, M.R.
 Method of producing thy a<->strains of vibrio cholerae, such
 strains and their use
 Patent: WO 9961634-A 1 02-DEC-1999;
 JOURNAL CARLIN NLS (SE); SBL VACCIN AB (SE); LEBENS MICHAEL R (SE)
 FEATRES
 Location/Qualifiers
 1..2909
 /organism="Vibrio cholerae"
 /mol_type="unassigned DNA"
 /db_xref="taxon:666"

ORIGIN

Query Match 100.0%; Score 2909; DB 6; Length 2909;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAGTTTGTATGCTCAGAGGTATCTGCAATTTCCCAATATTGACCCCGATTTGTT 60
DB 1 GAGAGTTTGTATGCTCAGAGGTATCTGCAATTTCCCAATATTGACCCCGATTTGTT 60
QY 61 TTGATCGGCGCTTACAGGAGGTGCGCTGATGAGTATTTGGTGGGTTCCCTTT 120
DB 61 TTGATCGGCGCTTACAGGAGGTGCGCTGATGAGTATTTGGTGGGTTCCCTTT 120
QY 121 TGTATGTTGGTCCCAATGCGCGAGCGGATGCGCGGAGTGTGGAAGCGGTGACA 180
DB 121 TGTATGTTGGTCCCAATGCGCGAGCGGATGCGCGGAGTGTGGAAGCGGTGACA 180
QY 181 AGCTGCTGATGTTATTTGCGCGGCTTTTAAAGTATGATCGGAGCGGATGTTA 240
DB 181 AGCTGCTGATGTTATTTGCGCGGCTTTTAAAGTATGATCGGAGCGGATGTTA 240
QY 241 TGTGATCTTCAATTTTATGATCTGTTCCCTGCTGACCTCTTTATTTTCAAGTGTG 300
DB 241 TGTGATCTTCAATTTTATGATCTGTTCCCTGCTGACCTCTTTATTTTCAAGTGTG 300
QY 301 GACTGCGGAGTATCTTCAAGGCGGCTTATTTGGTGTGATCAAGCGGATTTGCTGTA 360
DB 301 GACTGCGGAGTATCTTCAAGGCGGCTTATTTGGTGTGATCAAGCGGATTTGCTGTA 360
QY 361 TGGCGTAAACCAACGACCTTCTTGTGTGCGGATTTGTTGCTTATTTAGTGC 420
DB 361 TGGCGTAAACCAACGACCTTCTTGTGTGCGGATTTGTTGCTTATTTAGTGC 420
QY 421 ATTGCTTTGGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 480
DB 421 ATTGCTTTGGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 480
QY 481 AAGGATGTCCTTGGGCTTTTGTATTTCCCTAATGTTGCGGATGCGGATGCGGAT 540
DB 481 AAGGATGTCCTTGGGCTTTTGTATTTCCCTAATGTTGCGGATGCGGATGCGGAT 540
QY 541 AAGGATGTCCTTGGGCTTTTGTATTTCCCTAATGTTGCGGATGCGGATGCGGAT 600
DB 541 AAGGATGTCCTTGGGCTTTTGTATTTCCCTAATGTTGCGGATGCGGATGCGGAT 600
QY 601 TGTGTAACCTGCGGCTTACAGGAGGTGCGGATGCGGATGCGGATGCGGATGCGGAT 660
DB 601 TGTGTAACCTGCGGCTTACAGGAGGTGCGGATGCGGATGCGGATGCGGATGCGGAT 660
QY 661 CCGCTTCTTGTGGAATAGTCCGTGAGCGGATGCGGATGCGGATGCGGATGCGGAT 720
DB 661 CCGCTTCTTGTGGAATAGTCCGTGAGCGGATGCGGATGCGGATGCGGATGCGGAT 720
QY 721 CATTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 780
DB 721 CATTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 780
QY 781 TTGCTTTCAAGGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 840
DB 781 TTGCTTTCAAGGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 840
QY 841 GAAACAGATTTAGATTTTGTGACGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 900
DB 841 GAAACAGATTTAGATTTTGTGACGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 900
QY 901 ACGAAGCGGAGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 960
DB 901 ACGAAGCGGAGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 960
QY 961 CAATCAATTTCTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1020
DB 961 CAATCAATTTCTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1020
QY 1021 GCTGCGGATTTAGATTTTGTGACGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1080
DB 1021 GCTGCGGATTTAGATTTTGTGACGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1080
QY 1081 CTGGATGCTAATGCGGATTTTAAACCAAGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1140

DB 1081 CTGGATGCTAATGCGGATTTTAAACCAAGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1140
QY 1141 GATGACATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1200
DB 1141 GATGACATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1200
QY 1201 TCAATTTGACGATTTGAAAGATTTGATGATTTGACCGGATGCGGATGCGGATGCGGAT 1260
DB 1201 TCAATTTGACGATTTGAAAGATTTGATGATTTGACCGGATGCGGATGCGGATGCGGAT 1260
QY 1261 TGAATTTCTTAATCTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1320
DB 1261 TGAATTTCTTAATCTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1320
QY 1321 GTACAGGATCATTTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1380
DB 1321 GTACAGGATCATTTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1380
QY 1381 ATGTGATGCTCCCTTGGGATGAAATTTCAATGCGGATGCGGATGCGGATGCGGATGCGGAT 1440
DB 1381 ATGTGATGCTCCCTTGGGATGAAATTTCAATGCGGATGCGGATGCGGATGCGGATGCGGAT 1440
QY 1441 GATGCGGATCATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1500
DB 1441 GATGCGGATCATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1500
QY 1501 CATTTACCAAGATCACTGCAATGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1560
DB 1501 CATTTACCAAGATCACTGCAATGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1560
QY 1561 AGCGGCTCATGCTCATTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1620
DB 1561 AGCGGCTCATGCTCATTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1620
QY 1621 CATTTGATGATTTTACGCTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1680
DB 1621 CATTTGATGATTTTACGCTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1680
QY 1681 TTGATGCTAATCCCGATTTTCAAGGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1740
DB 1681 TTGATGCTAATCCCGATTTTCAAGGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1740
QY 1741 GGAAGGTCGAGGATTTTATTTATGATGATGCTTTTAAAGCGGATGCGGATGCGGATGCGGAT 1800
DB 1741 GGAAGGTCGAGGATTTTATTTATGATGATGCTTTTAAAGCGGATGCGGATGCGGATGCGGAT 1800
QY 1801 AATGCTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1860
DB 1801 AATGCTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1860
QY 1861 CTTACAGGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1920
DB 1861 CTTACAGGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1920
QY 1921 AATACCGTAAATCAAGGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1980
DB 1921 AATACCGTAAATCAAGGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1980
QY 1981 CAGAGCAACGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 2040
DB 1981 CAGAGCAACGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 2040
QY 2041 AGTGCAATGTTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 2100
DB 2041 AGTGCAATGTTTCAATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 2100
QY 2101 GCGGACGAGAGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 2160
DB 2101 GCGGACGAGAGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 2160
QY 2161 CACTGTACTGAAGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 2220

Db 2161 CACTGTACTGAAGCAATACCGTGTAGTGGGCTCGGCCAATCGCATTTTGTAGAAATTTC 2220

QY 2221 ACCGCGCGCGCCCAACCATCAAACTCTTCATCATGTTTCCCATCAACCGTAATGGCGAGAA 2280

Db 2221 ACCGCGCGCGCCCAACCATCAAACTCTTCATCATGTTTCCCATCAACCGTAATGGCGAGAA 2280

QY 2281 AATGTCGCAATACCCCAATAGTAAGTGGACACCCAGAAAGTATTAACCAATACCGA 2340

Db 2281 AATGTCGCAATACCCCAATAGTAAGTGGACACCCAGAAAGTATTAACCAATACCGA 2340

QY 2341 AAGCTGGGTTTCAAGCCCTGTGTATGACAGGTTTGGTAATGGTTGTATTAATAACCAACC 2400

Db 2341 AAGCTGGGTTTCAAGCCCTGTGTATGACAGGTTTGGTAATGGTTGTATTAATAACCAACC 2400

QY 2401 TTTCATCTCATATTCGCGCAGTCGCAAGCAGAGCGGAAAGACGACATGGAGCTTTTC 2460

Db 2401 TTTCATCTCATATTCGCGCAGTCGCAAGCAGAGCGGAAAGACGACATGGAGCTTTTC 2460

QY 2461 TAAATATCCCAACATCATTTCTAGAGTGAAGAAATCAGACCGCGAGAAATTTGAAAA 2520

Db 2461 TAAATATCCCAACATCATTTCTAGAGTGAAGAAATCAGACCGCGAGAAATTTGAAAA 2520

QY 2521 ATGTCGATGATGGGCACTGGCGAAAGCAGCGCCAAACTCTTTTACAGCGCATATGGCC 2580

Db 2521 ATGTCGATGATGGGCACTGGCGAAAGCAGCGCCAAACTCTTTTACAGCGCATATGGCC 2580

QY 2581 AAGGCTGACGAGAGTATTTGTCACAGTATGACCAATATTGGACCCATCAGCATAGAGAT 2640

Db 2581 AAGGCTGACGAGAGTATTTGTCACAGTATGACCAATATTGGACCCATCAGCATAGAGAT 2640

QY 2641 CGGCGTTTCAACCGGTATACCCACCGGCAACGAGACCAACATATATGAAGTACCGTGT 2700

Db 2641 CGGCGTTTCAACCGGTATACCCACCGGCAACGAGACCAACATATATGAAGTACCGTGT 2700

QY 2701 TGAAGATTTGATCAGTCCGCTTGCCACTAAACCATCATCAATCTCGAATTTGGGTGGGA 2760

Db 2701 TGAAGATTTGATCAGTCCGCTTGCCACTAAACCATCATCAATCTCGAATTTGGGTGGGA 2760

QY 2761 AGCAAAATTCAAATATGAACCTTTGCTGATGCGCGGTTGGCCATTAAAAACCGTGGCAG 2820

Db 2761 AGCAAAATTCAAATATGAACCTTTGCTGATGCGCGGTTGGCCATTAAAAACCGTGGCAG 2820

QY 2821 CATCGGACGTCGCAAGAGTAAATATACAGCATGAAAGCCAGTTTGGCCAAAGTAAAGCC 2880

Db 2821 CATCGGACGTCGCAAGAGTAAATATACAGCATGAAAGCCAGTTTGGCCAAAGTAAAGCC 2880

QY 2881 TTTCGTGATCAGCGAAATCGCGCTGCGAG 2909

Db 2881 TTTCGTGATCAGCGAAATCGCGCTGCGAG 2909

RESULT 3
LOCUS AE004153 9931 bp DNA linear BCT 10-OCT-2003
DEFINITION Vibrio cholerae O1 biovar eltor str. N16961 chromosome I, section 61 of 251 of the complete chromosome.
ACCESSION AE004153 AE003852
VERSION AE004153.1 GI:9655103
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

MEDLINE 20406833
PUBMED 10952301
REFERENCE
AUTHORS
2 (bases 1 to 9931)
Heidelberg, J.F., Eissen, J.A., Nelson, W.C., Clayton, R.A., Gwin, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.A.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ermlaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Uterback, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
Location/Qualifiers
1. 9931
/organism="Vibrio cholerae O1 biovar eltor str. N16961"
/mol_type="genomic DNA"
/strain="N16961"
/serotype="O1"
/db_xref="taxon:243277"
/chromosome="1"
/note="biotype: El Tor"
288. .602
/gene="VC0666"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAP93831.1"
/db_xref="GI:9655104"
/translation="VSSDFSIVIPSKTTGENTMRSIYLOLAVLVAAEAREER
VMKRVRRSSHYLPMNNPHLRDLETDGPRIGLSLPDAVVAERVRHRRVLAAT
PT"
complement(640. .1698)
/gene="VC0667"
complement(640. .1698)
/note="similar to PID:882727 GB:U00096 PID:1789199
PID:2342573; identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="oxidoreductase Tas, aldo/keto reductase family"
/protein_id="AAP93832.1"
/db_xref="GI:9655105"
/translation="MALRGSAMQYTKLPHSSLSISKICITMTFGHNSQAFOOL
DYALERSVNFIDTRKXVPVPTPTQOTGTEFIIGNMLAKSGKREKIVLATKVGPRV
PYLDRKALDRNTHQAVDSLRLOTDYIDLQHLWFORQTFPGQINPYDKOE
VTLLTLEALNDLVRMGKRVYIGVSNTEPWGWSYLRARKEHLPRTVSIQNYNLIN
RSFEVGLAEISHEGVGLAYSPPLAFGLGKYLNGARPAAGCTLHQRSEYFTEG
ILAEVAVLAQOFGDPAQMALAFVQRPFAVSNITGATTMEQLKSNLSDLSINA
ELLQIOEIGTTSNCP"
complement(1891. .2556)
/gene="VC0668"
complement(1891. .2556)
/note="similar to GB:Y00113 SP:P06722 GB:U16361 PID:42065
PID:577425; identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="DNA mismatch repair protein MucH"
/protein_id="AAP93833.1"
/db_xref="GI:9655106"
/translation="MKRAPPTQOELLTRAQIAGLSFAELADBAQMTVPDLKXGM
VGOLLEHLATAGSRPOCFEHLGIEIKSIPISTYKGLPETTFVCAAPLGVGHGTM
LGSVRAKSKVWIPQVGRERPLARCGSPGLMSPSEBAOLKADBEEMENIV
LGVAAQTLAKGSLVQLRPAANGRLTETVANGNRPKRLKRGVTLRQFTQIILOR
YXA"
complement(2943. .3035)
/gene="VC0669"
complement(2943. .3035)
/gene="VC0669"

Db 7254 AACGATGTCCTTGGGCTTTGTAATTCCTAATGATGAGCCACTGCCGCGCATCTTC 7313
 QY 541 AACGCTTATGAAATGCGCTTAGAAGCGGTGTTCTGTTCTTATCTTAATGTTAT 600
 Db 7314 AACGCTTATGAAATGCGCTTAGAAGCGGTGTTCTTATCTTAATGTTAT 7373
 QY 601 TGGTAAACCTGTCGCTAGGAGCGATCCGAGCTGTTTATAGTGAATAGGTACAT 660
 Db 7374 TGGTAAACCTGTCGCTAGGAGCGATCCGAGCTGTTTATAGTGAATAGGTACAT 7433
 QY 661 CCGCTTCTTGTGAAATAGCTCCGTAGAGCGAATGCTAATGGTGTGTTGTTGCTT 720
 Db 7434 CCGCTTCTTGTGAAATAGCTCCGTAGAGCGAATGCTAATGGTGTGTTGCTT 7493
 QY 721 CATTTAAATGGGGCAATCCTCCTTACTTAATGATGATCATGATATTTGATATGAT 780
 Db 7494 CATTTAAATGGGGCAATCCTCCTTACTTAATGATGATCATGATATTTGATATGAT 7553
 QY 781 TTGGTCTTACAAAGCGCGGTTTGTATCAGACCGTGTAGAGCAAAATAGGTAAGT 840
 Db 7554 TTGGTCTTACAAAGCGCGGTTTGTATCAGACCGTGTAGAGCAAAATAGGTAAGT 7613
 QY 841 GAAACGATATTGATCTTTGTAGCGCATCGTCAATCAAAGTGTGTTGTTGAAATGA 900
 Db 7614 GAAACGATATTGATCTTTGTAGCGCATCGTCAATCAAAGTGTGTTGTTGAAATGA 7673
 QY 901 ACGAAGCGGCAAGCGTGTGTGACTGTGATTAATGCGAATTTGACCTACATGTTGGGCA 960
 Db 7674 ACGAAGCGGCAAGCGTGTGTGACTGTGATTAATGCGAATTTGACCTACATGTTGGGCA 7733
 QY 961 CAATCATGTTTCTTATGATCAACGCAAGTTTGTGAAAGCTGCCGTAGCGAGT 1020
 Db 7734 CAATCATGTTTCTTATGATCAACGCAAGTTTGTGAAAGCTGCCGTAGCGAGT 7793
 QY 1021 GCTCGGCTATATTCGCTGTAGCAATATGCGGCGATTTTCCGCAATTTAGTACCAAAAC 1080
 Db 7794 GCTCGGCTATATTCGCTGTAGCAATATGCGGCGATTTTCCGCAATTTAGTACCAAAAC 7853
 QY 1081 CTGGGATGCTAAATGCCAATTTAAACCAAGCATGCTCAACAATCTTACCTTAAGTGA 1140
 Db 7854 CTGGGATGCTAAATGCCAATTTAAACCAAGCATGCTCAACAATCTTACCTTAAGTGA 7913
 QY 1141 GATGATCATGGAAGCGTGTATGTTGTAGAGTGAAGTTGGGCTTAAGCTGATGTGG 1200
 Db 7914 GATGATCATGGAAGCGTGTATGTTGTAGAGTGAAGTTGGGCTTAAGCTGATGTGG 7973
 QY 1201 TCATATTTGACCAAGTTGAAAAGATTGTGATGATTTGAGCCGTGGCTGATGACGAG 1260
 Db 7974 TCATATTTGACCAAGTTGAAAAGATTGTGATGATTTGAGCCGTGGCTGATGACGAG 8033
 QY 1261 TGAATTTCTTAATCTTCAATCCGAGTGAATTTCAATGAGGTTGTTGCCGCTTGAT 1320
 Db 8034 TGAATTTCTTAATCTTCAATCCGAGTGAATTTCAATGAGGTTGTTGCCGCTTGAT 8093
 QY 1321 GTACAGCCATCAATTTTCAATGCTGGGGATACCTTGAATCTCAACATGATCAGCGTTC 1380
 Db 8094 GTACAGCCATCAATTTTCAATGCTGGGGATACCTTGAATCTCAACATGATCAGCGTTC 8153
 QY 1381 ATGTATGTGCTTTGGGTTGAATTTCAATGATGAGGTTTATGTTGCTTCTGCGCT 1440
 Db 8154 ATGTATGTGCTTTGGGTTGAATTTCAATGATGAGGTTTATGTTGCTTCTGCGCT 8213
 QY 1441 GATGACAGATCAACAGGAAAAGCCGGGCTTGGCGTATCAACAATGCTGATGCGCA 1500
 Db 8214 GATGACAGATCAACAGGAAAAGCCGGGCTTGGCGTATCAACAATGCTGATGCGCA 8273
 QY 1501 CATTTACCAAGATCAATCGAATGATGCGGATGCGATTAACCTGAGCCATTGCC 1560
 Db 8274 CATTTACCAAGATCAATCGAATGATGCGGATGCGATTAACCTGAGCCATTGCC 8333
 QY 1561 AGCGGCTCAATTTCAATCAATCCAAAGTTTAAACCTGAGGATTTGAAACTTGCT 1620
 Db 8334 AGCGGCTCAATTTCAATCAATCCAAAGTTTAAACCTGAGGATTTGAAACTTGCT 8393

QY 1621 CACTTGTGATGATTTGACGTCACCGGATATCAGTTCCACAGATCTATTAATCCGCTT 1680
 Db 8394 CACTTGTGATGATTTGACGTCACCGGATATCAGTTCCACAGATCTATTAATCCGCTT 8453
 QY 1681 TTGAGCTTAATCCGCTATTTCAAGGGGTATGCTTGTATGATGATTTTATATTAAGTCTC 1740
 Db 8454 TTGAGCTTAATCCGCTATTTCAAGGGGTATGCTTGTATGATGATTTTATATTAAGTCTC 8513
 QY 1741 CGAAGTCCGAGGCTTTTATATACAGATGATGCTTTAACCTTAAGGGTTAGGGCAAG 1800
 Db 8514 CGAAGTCCGAGGCTTTTATATACAGATGATGCTTTAACCTTAAGGGTTAGGGCAAG 8573
 QY 1801 AATGTCGCGGGGATGACGACAAACACCCATTAAGTAACTACCAACCACTTTGCT 1860
 Db 8574 AATGTCGCGGGGATGACGACAAACACCCATTAAGTAACTACCAACCACTTTGCT 8633
 QY 1861 CTTACAAAGCCCAAGTGTAGATGAGCTCAGACCTTTAATGAGGCTTGGGTAGAAAG 1920
 Db 8634 CTTACAAAGCCCAAGTGTAGATGAGCTCAGACCTTTAATGAGGCTTGGGTAGAAAG 8693
 QY 1921 AATACTGTAATCAAGACCGTATGACATCAAGTTAAAGCTTAAGTCAACACCGCAATTTG 1980
 Db 8694 AATACTGTAATCAAGACCGTATGACATCAAGTTAAAGCTTAAGTCAACACCGCAATTTG 8753
 QY 1981 CAGAGCAAAACCGGCAAACTCACAGAGACAGCGGTTGGCGGACGAGACGATTAACA 2040
 Db 8754 CAGAGCAAAACCGGCAAACTCACAGAGACAGCGGTTGGCGGACGAGACGATTAACA 8813
 QY 2041 AGTCCATGTTCCACACTAAGTAAGTAATGATTTCAAGCACTTTCAGACGCGAGA 2100
 Db 8814 AGTCCATGTTCCACACTAAGTAATGATTTCAAGCACTTTCAGACGCGAGA 8873
 QY 2101 GCCCAGAGAGAAACCTTATGCTGTTGTGTGATGAATGAATGAATTAATACGTAAC 2160
 Db 8874 GCCCAGAGAGAAACCTTATGCTGTTGTGTGATGAATGAATGAATTAATACGTAAC 8933
 QY 2161 CACTGTACCTGAAGCAATTAACGTTATGAGGCTTGGCGCAATGCAATTTGTAATTTG 2220
 Db 8934 CACTGTACCTGAAGCAATTAACGTTATGAGGCTTGGCGCAATGCAATTTGTAATTTG 8993
 QY 2221 ACGTCCGCGCAACATCAAACTCTTCAATCAAGTTTCCCATCAACCGTAATGCGAGAA 2280
 Db 8994 ACGTCCGCGCAACATCAAACTCTTCAATCAAGTTTCCCATCAACCGTAATGCGAGAA 9053
 QY 2281 AATGTCGCAATACCCCAATAGATGAATGAGCAACCAACCGAATTAACCAATACCGA 2340
 Db 9054 AATGTCGCAATACCCCAATAGATGAATGAGCAACCAACCGAATTAACCAATACCGA 9113
 QY 2341 AAGCTGGGTTTCAAGCCCTGTGATGACAGGTTGTGTAATGCTTTGTAATAACCAAAAC 2400
 Db 9114 AAGCTGGGTTTCAAGCCCTGTGATGACAGGTTGTGTAATGCTTTGTAATAACCAAAAC 9173
 QY 2401 TTTCAATGCTCATATCGCAATCGCAAGCAAGCGGAAACGACGATGTGAGCTTCTC 2460
 Db 9174 TTTCAATGCTCATATCGCAATCGCAAGCAAGCGGAAACGACGATGTGAGCTTCTC 9233
 QY 2461 TAAATGCAAAATCATTTCTAGAGTGAAGATGACAGCGGCAAAAGTTGAAAAA 2520
 Db 9234 TAAATGCAAAATCATTTCTAGAGTGAAGATGACAGCGGCAAAAGTTGAAAAA 9293
 QY 2521 ATGTGATGTGTGCACTGCGAAGCAAGCGGCAAACTCTTCTTAAACAGGCAATGTGCT 2580
 Db 9294 ATGTGATGTGTGCACTGCGAAGCAAGCGGCAAACTCTTCTTAAACAGGCAATGTGCT 9353
 QY 2581 AAGGTCGACAGATGTTGCTCAAGTATGATCAATTAATGAGCAAGCAATCAATAGAAAT 2640
 Db 9354 AAGGTCGACAGATGTTGCTCAAGTATGATCAATTAATGAGCAAGCAATCAATAGAAAT 9413
 QY 2641 CGCGGTTTCAACCGGTAACCCACCGGCAACGAGACCAACATTAATGAAGTCAACGCTGCT 2700
 Db 9414 CGCGGTTTCAACCGGTAACCCACCGGCAACGAGACCAACATTAATGAAGTCAACGCTGCT 9473

QY 2701 TGAGGATTGAATCACTGCGCTTGGCACTAAACCAATCATCATCTGCAATTGGGTGGGA 2760
DB 9474 TGAGGATTGAATCACTGCGCTTGGCACTAAACCAATCATCATCTGCAATTGGGTGGGA 9533
QY 2761 AGCAATTCATTAATGAATCTTGGCTTGAATGCGCGCTTGGCACTAAACCGTGGCGAC 2820
DB 9534 AGCAATTCATTAATGAATCTTGGCTTGAATGCGCGCTTGGCACTAAACCGTGGCGAC 9593
QY 2821 CATCGGCACTGCAAGAACTAGTAATAACAGCATGAAAGCAAGTTGGCCCAAGTGGCC 2880
DB 9594 CATCGGCACTGCAAGAACTAGTAATAACAGCATGAAAGCAAGTTGGCCCAAGTGGCC 9653
QY 2881 TTTCGTGTCAGCGCAATCGCGCTGCGAG 2909
DB 9654 TTTCGTGTCAGCGCAATCGCGCTGCGAG 9682

RESULT 4
VCTHYAG 1410 bp DNA linear BCT 01-NOV-2000
LOCUS Vibrio cholerae thya gene.
DEFINITION Y17135
ACCESSION Y17135
VERSION Y17135.1 GI:3114980
KEYWORDS thya gene; thymidilate synthase.
SOURCE Vibrio cholerae
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.

REFERENCE 1
AUTHORS Valle, E., Ledon, T., Cedre, B., Campos, J., Valmaseda, T., Rodriguez, B., Garcia, L., Marrero, K., Benitez, J., Rodriguez, S. and Fando, R.
TITLE Construction and characterization of a nonproliferative El Tor cholera vaccine candidate derived from strain 638
JOURNAL Infect. Immun. 68 (11), 6411-6418 (2000)
MEDLINE 20490577
PUBMED 11035753
REFERENCE 2 (bases 1 to 1410)
AUTHORS Benitez, J.A.
TITLE Direct Submission
JOURNAL Submitted (16-APR-1998) J.A. Benitez, Centro Nacional de Investigaciones, Cientificas, PO Box 6990, La Habana, CUBA

FEATURES
source
1..1410
/organism="Vibrio cholerae"
/mol_type="genomic DNA"
/strain="C7258"
/db_xref="taxon:666"
289..1140
/gene="thya"
289..1140
/gene="thya"
/codon_start=1
/transf_table=11
/product="thymidilate synthase"
/protein_id="CAA6645.1"
/db_xref="GI:3114981"
/db_xref="GOA:O66108"
/db_xref="SWISS-PROT:O66108"
/translation="MROYLDCORIVDOGWEMENETGRKCLTVINADLTVDVGNNOFPIVTRKSPKAAVAELIGYIRGDNAADPRLGCTRTWNNANINPCANLPPYRGEDDMGRVYGVGRAMAKPDGHIIDDLKIVDDLSGDDREIILNPNPBYHMGCLRPCMYSHFSLDGTIFLNSDKSCVPGLNPNVQYVPLALPAQITGKPGLAHKIVNAHYDQLELMRDVQIKREPFAPQFHNPKIKTLQDLELWVLDDPDTGYGFHPDIQYFESV"

ORIGIN
Query Match 47.8%; Score 1390.8; DB 1; Length 1410;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1398; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 551 GAATTCGCTTGAAGAGCGGTGCTTCTTATTCTTAATGTTATTTGTAACCT 610
|||||
DB 1 GAATTCGCTTGAAGAGCGGTGCTTCTTATTCTTAATGTTATTTGTAACCT 610

DB 1 GAATTCGCTTGAAGAGCGGTGCTTCTTATTCTTAATGTTATTTGTAACCT 60
QY 611 CGTCGCTAGACAGCGGATCCGACGTGTTTATAGTGATACGATACATCCGTCCTT 670
DB 61 CGTCGCTAGACAGCGGATCCGACGTGTTTATAGTGATACGATACATCCGTCCTT 120
QY 671 GTGGAATAGCTCCGTGAGCCAGATGCTCACTGGGTCTGTTGGTGGCTTCAATTCATG 730
DB 121 GTGGAATAGCTCCGTGAGCCAGATGCTCACTGGGTCTGTTGGTGGCTTCAATTCATG 180
QY 731 GGGCAATTCCTCTCTTACCTATGATGATCATGGATTTTATGATAGTGTGCTTAC 790
DB 181 GGGCAATTCCTCTCTTACCTATGATGATCATGGATTTTATGATAGTGTGCTTAC 240
QY 791 AAGCGCGTTTGTATCAAGCGGTGAGCAAGCAAAATAGGATAGTTAGTGAACATAT 850
DB 241 AAGCGCGTTTGTATCAAGCGGTGAGCAAGCAAAATAGGATAGTTAGTGAACATAT 300
QY 851 TTGATCTTTGTCAGGCAATCGTCATCAAGGTGTTGGTGAATAAGCAAGGAG 910
DB 301 TTGATCTTTGTCAGGCAATCGTCATCAAGGTGTTGGTGAATAAGCAAGGAG 360
QY 911 AAGCGTTTGTACTGTGATTAATGCCATTTGACCTACAGATGAGGCAACATAGTT 970
DB 361 AAGCGTTTGTACTGTGATTAATGCCATTTGACCTACAGATGAGGCAACATAGTT 420
QY 971 CCTCTAGTACTACGCAAGAGATTGTTGAAGCTGCGTAGCGAGTTGCTGGCTAT 1030
DB 1031 ATTGCTGTTAGCATTAATGCGGAGATTTTGGCCATTAAGTACCAAACTGGGATGCT 1090
QY 481 ATTGCTGTTAGCATTAATGCGGAGATTTTGGCCATTAAGTACCAAACTGGGATGCT 540
DB 1091 AATGCCATTTAAACCAAGCATGCTGCAACATCTCTCCGTAAGGAGATGACATG 1150
QY 541 AATGCCATTTAAACCAAGCATGCTGCAACATCTCTCCGTAAGGAGATGACATG 600
DB 1151 GAGCGGTGATGCTGTTAGGCTGAGCTTGGCTTAAGCTGATGCTGCTCATTTGAC 1210
QY 601 GAGCGGTGATGCTGTTAGGCTGAGCTTGGCTTAAGCTGATGCTGCTCATTTGAC 660
DB 1211 CAGTTGAAAAGATGTTGATGATTTGAGCGCTGCTGATGACCGAGGTGAATGTT 1270
QY 661 CAGTTGAAAAGATGTTGATGATTTGAGCGCTGCTGATGACCGAGGTGAATGTT 720
DB 1271 AACTTCAATCCGCGGTAAATTTCAATGCGGTGTTTGGCGCTTGCATGATGACCAT 1330
QY 721 AACTTCAATCCGCGGTAAATTTCAATGCGGTGTTTGGCGCTTGCATGATGACCAT 780
DB 1331 CATTTTCAATGCTGCGGTAAATTTCAATGCGGTGTTTGGCGCTTGCATGATGATG 1390
QY 781 CATTTTCAATGCTGCGGTAAATTTCAATGCGGTGTTTGGCGCTTGCATGATGATG 840
DB 1391 CCTTGGGTTGAATTTCAACATGCTGCAAGTTATGTTGCTTGGCTGATGACGAC 1450
QY 841 CCTTGGGTTGAATTTCAACATGCTGCAAGTTATGTTGCTTGGCTGATGACGAC 900
DB 1451 ATCACAGGGAAGAGCGGCTTGGCGTATCAAGATCGTCAATGGGCAATTTACCA 1510
QY 901 ATCACAGGGAAGAGCGGCTTGGCGTATCAAGATCGTCAATGGGCAATTTACCA 960
DB 1511 GATCAATCGAATGATGCGGATGCTGCAAGCTTAAAGCGATTTCCAGCGCTGAG 1570
QY 961 GATCAATCGAATGATGCGGATGCTGCAAGCTTAAAGCGATTTCCAGCGCTGAG 1020
DB 1571 TTCATATCAATCAAGATTTAAACACGTGCGAGTTTGAACCTGGGTCACTTTGAT 1630
QY 1021 TTCATATCAATCAAGATTTAAACACGTGCGAGTTTGAACCTGGGTCACTTTGAT 1080
DB 1631 GATTTGACGTACCGGATATCATGTTCAACATCTTATCAATACCGGTTTCACTTAA 1690
QY 1081 GATTTGACGTACCGGATATCATGTTCAACATCTTATCAATACCGGTTTCACTTAA 1140

QY 1691 TCCGATTCAGGCGGATGAGTGGCTGATNGGCTTTTATATATAAAAGTCCCGAAGTCCG 1750
DB 1141 TCCGATTCAGGCGGATGAGTGGCTGATNGGCTTTTATATATAAAAGTCCCGAAGTCCG 1200
QY 1751 GAGCTTTTATATACAGATGATGCTTTAAGCCTTAAGCGGTTAAGGCGAAGATGTCGCCG 1810
DB 1201 GAGCTTTTATATACAGATGATGCTTTAAGCCTTAAGCGGTTAAGGCGAAGATGTCGCCG 1260
QY 1811 GGAATGACGCAAAACACACACCCCAATAGTAATCAACACCACTTTGCTTTACAGCC 1870
DB 1261 GGAATGACGCAAAACACACACCCCAATAGTAATCAACACCACTTTGCTTTACAGCC 1320
QY 1871 CAAGTTGAGATGAGTCAAGCCTTTAATAGGAGTTCGGTAAAGGAAGGAATACCGTAA 1930
DB 1321 CAAGTTGAGATGAGTCAAGCCTTTAATAGGAGTTCGGTAAAGGAAGGAATACCGTAA 1380
QY 1931 ATCAAGACCGTAGCCATCAAGTTAAAGCTT 1960
DB 1381 ATCAAGACCGTAGCCATCAAGTTAAAGCTT 1410

RESULT 5
AX009483
LOCUS AX009483 1222 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 3 from Patent WO961634.
ACCESSION AX009483
VERSION AX009483.1 GI:9996768
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae
Vibrio cholerae
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.

REFERENCE
1 Carlin N. and Lebens M.R.
AUTHORS Method of producing thy a<->strains of vibrio cholerae, such
TITLE strains and their use
JOURNAL Patent: WO 961634-A 3 02-DEC-1999;
CARLIN NILS (SE); SBL VACCIN AB (SE); LEBENS MICHAEL R (SE)
FEATURES
source
1. 1222
/organism="Vibrio cholerae"
/mol_type="unassigned DNA"
/db_xref="taxon:666"

ORIGIN
Query Match 42.0%; Score 1222; DB 6; Length 1222;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1688 TAAATCCCGATTCAGGCGGATGAGTGGCTTTTATATATAAAAGTCCCGAAGT 1747
DB 1 TAAATCCCGATTCAGGCGGATGAGTGGCTTTTATATATAAAAGTCCCGAAGT 60
QY 1748 CCGGAGCTTTTATACAGATGATGCTTTAAGCCTTAAGCGGTTAAGGCGAAGATGCTG 1807
DB 61 CCGGAGCTTTTATACAGATGATGCTTTAAGCCTTAAGCGGTTAAGGCGAAGATGCTG 120
QY 1808 CCGGAGATGACAAACACACACCCCAATAGTAATCAACACCACTTTGCTTTACAA 1867
DB 121 CCGGAGATGACAAACACACACCCCAATAGTAATCAACACCACTTTGCTTTACAA 180
QY 1868 GCCCAAGTTGAGATGAGTCAAGCCTTTAATAGCAATTTGGCTTAAGGAATACCG 1927
DB 181 GCCCAAGTTGAGATGAGTCAAGCCTTTAATAGCAATTTGGCTTAAGGAATACCG 240
QY 1928 TAAATCAAGACCGTAGCCATCAAGTTAAAGCTTAAGTCAACACGCAATTTGCAAGCA 1987
DB 241 TAAATCAAGACCGTAGCCATCAAGTTAAAGCTTAAGTCAACACGCAATTTGCAAGCA 300
QY 1988 AACACGGCAAACTCACCGAGACACCGGTTGGGCGACAGACAGCATATCAAGTGCCA 2047
DB 301 AACACGGCAAACTCACCGAGACACCGGTTGGGCGACAGACAGCATATCAAGTGCCA 360

QY 2048 ATGTTGCACTTAAGTAATAGGTAGATTTCACGCACTTTACAGACCCAGAGCCACG 2107
DB 361 ATGTTGCACTTAAGTAATAGGTAGATTTCACGCACTTTACAGACCCAGAGCCACG 420
QY 2108 AGAGGAACCAATAGGCTGTTGTGTCATGAAGATTAACTAATACCGTAAACCACTGTA 2167
DB 421 AGAGGAACCAATAGGCTGTTGTGTCATGAAGATTAACTAATACCGTAAACCACTGTA 480
QY 2168 CCTGAAGCAATACCGTGTAGTGGGCTTGCCCAATTCGATTTTGTGAATTTCAAGTGC 2227
DB 481 CCTGAAGCAATACCGTGTAGTGGGCTTGCCCAATTCGATTTTGTGAATTTCAAGTGC 540
QY 2228 CGGCAACCAATCAACTTTTATCATGATTTGCCCATACCGTAAATGCGACGAAATGTC 2287
DB 541 CGGCAACCAATCAACTTTTATCATGATTTGCCCATACCGTAAATGCGACGAAATGTC 600
QY 2288 GCAATACCAATACGATTAAGTGCACACACCGAAAGTATTACCAATACCGAAAGCTGC 2347
DB 601 GCAATACCAATACGATTAAGTGCACACACCGAAAGTATTACCAATACCGAAAGCTGC 660
QY 2348 GTTTCAGCCCTGTGATGACAGGTTTGTGTAATTCGTTTGAATTAATCAAACTTTTCATG 2407
DB 661 GTTTCAGCCCTGTGATGACAGGTTTGTGTAATTCGTTTGAATTAATCAAACTTTTCATG 720
QY 2408 CTCATATGCGCAGTCCGACAGAGAGGCGAAACGAGCCAGTGTGACTTTCTTAAATG 2467
DB 721 CTCATATGCGCAGTCCGACAGAGAGGCGAAACGAGCCAGTGTGACTTTCTTAAATG 780
QY 2468 CCAACATCATTTCTTAAGTAGTAAGTAAGTACGACCCGCGAAGATGAAAAATCGTG 2527
DB 781 CCAACATCATTTCTTAAGTAGTAAGTAAGTACGACCCGCGAAGATGAAAAATCGTG 840
QY 2528 ATGTGGCATCTGGGGAAGACAGCGCGCAACTCTTTTAAAGCGCATATGCGCAAGCTG 2587
DB 841 ATGTGGCATCTGGGGAAGACAGCGCGCAACTCTTTTAAAGCGCATATGCGCAAGCTG 900
QY 2588 ACGAGATATTGTCACAGTAGTACCATATTGCGCACTTCCATACCATAGGAATGCGGTT 2647
DB 901 ACGAGATATTGTCACAGTAGTACCATATTGCGCACTTCCATACCATAGGAATGCGGTT 960
QY 2648 TCAACCGGTAAACCCAGCGGACGACGACCAACATTAATAGATCAACCGTCTTGAGAT 2707
DB 961 TCAACCGGTAAACCCAGCGGACGACGACCAACATTAATAGATCAACCGTCTTGAGAT 1020
QY 2708 TGAATCAGTCCGCTGTCACATTAACCAATCATCAATCTTGTGTTGGTGGGAAACAAAT 2767
DB 1021 TGAATCAGTCCGCTGTCACATTAACCAATCATCAATCTTGTGTTGGTGGGAAACAAAT 1080
QY 2768 TCAATTAAGAACTTTGCTTATGCGCGGTTGCCATTTAAACCGGTCGCCAATCGCG 2827
DB 1081 TCAATTAAGAACTTTGCTTATGCGCGGTTGCCATTTAAACCGGTCGCCAATCGCG 1140
QY 2828 ACTGCAAGATGTAAATACAGATGAAGCCAAAGTTTCCCAAGTAAAGCTTTTCGTG 2887
DB 1141 ACTGCAAGATGTAAATACAGATGAAGCCAAAGTTTCCCAAGTAAAGCTTTTCGTG 1200
QY 2888 GTCAGCGAAATCGGCGTGCAG 2909
DB 1201 GTCAGCGAAATCGGCGTGCAG 1222

RESULT 6
AY143429
LOCUS AY143429 1236 bp DNA linear BCT 28-SEP-2002
DEFINITION Vibrio cholerae mutant thymidilate synthetase (thyA) gene, complete
cds.
ACCESSION AY143429
VERSION AY143429.1 GI:23343945
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae
Vibrio cholerae
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

REFERENCE 1 (bases 1 to 1236)
 AUTHORS Xia X.
 TITLE Direct Submission
 JOURNAL Submitted (20-AUG-2002) Microbiology, Institute for Infectious Disease Control and Prevention, Chinese Center for Disease Control and Prevention, P.O. Box 5, Changping, Beijing 102206, P.R. China

FEATURES
 source 1..1236
 /organism="Vibrio cholerae"
 /mol_type="genomic DNA"
 /db_xref="taxon:566"

gene

CDS

248..1099
 /gene="thyA"
 /note="thyA"
 /gene="thyA"
 /note="thyA"
 /note="does not catalyze the conversion of dUTP into dTMP"
 /codon_start=1
 /transl_table=11
 /product="mutant thymidilate synthetase"
 /protein_id="AA016904.1"
 /db_xref="GI:23343946"
 /translation="MQOYLDLCRIVDQGWVNERGKCVINADLYDVGNNGP
 PLVTRKSPFMAVAELGYIRGDNADPQLGTGTWDANANLNQANNPYRKGED
 DMGRVYVCGGRAMAKPDGGHIDOLKIVDLSRGVDDGEILNFPNPEPMGCLP
 MYSHPSLIGDTLYLNSRSCDPLMINNMVVOYVFLALMAOITGKPGIAYKIV
 NHHYDDJLMDVQLKREPPAPQFHNPKIKLQDLFTWVTLDDSDVGYGFHD
 IQYPSV"

ORIGIN

Query Match 41.8%; Score 1216.8; DB 1; Length 1236;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 1224; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 592 TTGATTATGTAACCTGCTCCGCTAGGACGCGTATCCGACATGTTTATGCTGGATA 651
 DB 1 TTGATTATGTAACCTGCTCCGCTAGGACGCGTATCCGACATGTTTATGCTGGATA 60
 QY 652 CGGTACATTCGGCTTCCTGTGGAATACGTCGAGCAGATGCTCAGTTGGTCTGTT 711
 DB 61 CGGTACATTCGGCTTCCTGTGGAATACGTCGAGCAGATGCTCAGTTGGTCTGTT 120
 QY 712 TGGTGGCTTCATTCATGAGGCGCAATCCTCTCTTACCTATGATGATCATGCTATTTT 771
 DB 121 TGGTGGCTTCATTCATGAGGCGCAATCCTCTCTTACCTATGATGATCATGCTATTTT 180
 QY 772 GATGATGCTTGTGCTTCAACAGCGCGTGTGATCAAGCCGTGACGCAAAATATAGG 831
 DB 181 GATGATGCTTGTGCTTCAACAGCGCGTGTGATCAAGCCGTGACGCAAAATATAGG 240
 QY 832 TAGTTAGTGAACAGATATTAGATCTTGTGTCAGCGCATGTCGATCAAAGTGTGGGT 891
 DB 241 TAGTTAGTGAACAGATATTAGATCTTGTGTCAGCGCATGTCGATCAAAGTGTGGGT 300
 QY 892 TGAATATGAACGAGCGGCAAGCGTGTGATGATGATTAAGCCGATTTGACCTACGA 951
 DB 301 TGAATATGAACGAGCGGCAAGCGTGTGATGATGATTAAGCCGATTTGACCTACGA 360
 QY 952 TGTGGGCAACAATAGTTTCTCTAGTACATACGCAAGAGTTTGGAAAGCTCCGT 1011
 DB 361 TGTGGGCAACAATAGTTTCTCTAGTACATACGCAAGAGTTTGGAAAGCTCCGT 420
 QY 1012 AGCCGAGTGTGCTGGCTATATTTGTGTTACGATATGCGCGGATTTTGCCATTAAG 1071
 DB 421 AGCCGAGTGTGCTGGCTATATTTGTGTTACGATATGCGCGGATTTTGCCATTAAG 480
 QY 1072 TACCAAAACCTGGGATGCTATAGCCATTTAAACCAAGCATGGCTCAACATCCTTACCG 1131
 DB 481 TACCAAAACCTGGGATGCTATAGCCATTTAAACCAAGCATGGCTCAACATCCTTACCG 540
 QY 1132 TAAAGTGAGATGACATGAGGACGCGTGTATGAGATTGAGGATAGAGCTTGGGCTAAGCC 1191

DB 541 TAAAGTGAGATGACATGAGGACGCGTGTATGAGATTGAGGATAGAGCTTGGGCTAAGCC 600
 QY 1192 TGAATGCTGCTATGATGACAGTGAAGAAAGATTGTTGATGATTTAGCCGTGCGCTTGA 1251
 DB 601 TGAATGCTGCTATGATGACAGTGAAGAAAGATTGTTGATGATTTAGCCGTGCGCTTGA 660
 QY 1252 TGACCGAGTGAAATCTTAACTTCTTCAACATCCGGGTGAATTTACATGAGGCTTTGGG 1311
 DB 661 TGACCGAGTGAAATCTTAACTTCTTCAACATCCGGGTGAATTTACATGAGGCTTTGGG 720
 QY 1312 CCTTGCATGTAAGGCATCATTTTTCATGCTGGGGGATPACTGTTATCTCAACAGTAC 1371
 DB 721 CCTTGCATGTAAGGCATCATTTTTCATGCTGGGGTGAATCTTATCTCAACAGTAC 780
 QY 1372 TCAGCGTCAATGATGTCGCTTGGGGTGAATTTCAACATGTCGAGCTTTATGCTTT 1431
 DB 781 TCAGCGTCAATGATGTCGCTTGGGGTGAATTTCAACATGTCGAGCTTTATGCTTT 840
 QY 1432 CTTTGGCTGATGAGCAGATCACTCAGGAAAAAGCCGGGCTTGGGATCAACAAGATCGT 1491
 DB 841 CTTTGGCTGATGAGCAGATCACTCAGGAAAAAGCCGGGCTTGGGATCAACAAGATCGT 900
 QY 1492 CAATGGCAGATTACCAAGATCACTCAGATGATGCGCGATGTCAGCTTAAACGTGA 1551
 DB 901 CAATGGCAGATTACCAAGATCACTCAGATGATGCGCGATGTCAGCTTAAACGTGA 960
 QY 1552 GCCATTCCAGCGCCTCAGTTCATATCAATCCAAAGATTAAACCTGACAGATTTGGA 1611
 DB 961 GCCATTCCAGCGCCTCAGTTCATATCAATCCAAAGATTAAACCTGACAGATTTGGA 1020
 QY 1612 AACTGGCTACATTGGATGATTTTGAACGTCACCGATATCACTGTCAGATCCATATTA 1671
 DB 1021 AACTGGCTACATTGGATGATTTTGAACGTCACCGATATCACTGTCAGATCCATATTA 1080
 QY 1672 ATACCCGTTTCACTGATTAATCCCGTATTCAGCGGATGAGCTTGAATGCTTTATATTA 1731
 DB 1081 ATACCCGTTTCACTGATTAATCCCGTATTCAGCGGATGAGCTTGAATGCTTTATATTA 1140
 QY 1732 AAAGCTCCGGAAGTGGAGCTTTTATATACAGTATGCTTAAAGCTTAAAGCGGT 1791
 DB 1141 AAAGCTCCGGAAGTGGAGCTTTTATATACAGTATGCTTAAAGCTTAAAGCGGT 1200
 QY 1792 TAGGCAAGATGCTGCGGCGATGACGACAAACAC 1827
 DB 1201 TAGGCAAGATGCTGCGGCGATGACGACAAACAC 1236

RESULT 7
 VCAJ10968/c 1591 bp DNA linear BCT 21-SEP-1998
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 Vibrionaceae; Vibrio.
 REFERENCE
 1 Lebens M., Soderlund L.O., Lundquist P. and Carlin N.I.A.
 A putative sodium dependent phosphate pump in Vibrio cholerae
 Unpublished
 JOURNAL
 TITLE
 AUTHORS
 REFERENCE
 2 (bases 1 to 1591)
 Carlin N.I.A.
 Direct Submission
 Submitted (17-SEP-1998) Carlin N.I.A., SBL Vaccin AB, Department of
 Molecular biology, FMH, UV 30, Stockholm, S-105 21, SWEDEN
 npta gene is located between the nhar gene (AU002395) and the
 thyA gene (AU006514) both running in opposite directions.
 COMMENT
 FEATURES
 source 1..1591
 /organism="Vibrio cholerae"

```

/mol_type="genomic DNA"
/strain="US1569"
/db_xref="taxon:666"
245..1588
-gene="nptA"
-35_signal
245..250
-gene="nptA"
-10_signal
267..272
-gene="nptA"
373..376
-gene="nptA"
383..1532
-gene="nptA"
/codon_start=1
/translation="sodium dependent phosphate pump"
/product="sodium dependent phosphate pump"
/protein_id="CAA09443.1"
/db_xref="GI:3646477"
/db_xref="GOA:O87918"
/db_xref="SPTREMBL:O87918"
/translation="MINOTSPAPISLTTKGLRWANLAFMLYLLLAAMVSGFKX
ATGDAKLVFEFASHPIAGIMGLVATLIGSSVTSTIIIVLAVGGIPVETALPMV
GANIGTTVNTLIVSGHRCKEFRAPASATIHDPNLLATILPLPMMFGILERY
SHWISPLLATGDMKSKFPIKPIITKPVITGLTSLTGLTGVGLIVGLATIF
VAITVGLKMLKLMVGRAREILQNAIGRPGLIASGVTVTVLVOSSTTSLMVPV
GSGVLRREIYPTLGTAGNIGTITALLATATVGEFAVFAIQIALVHLSPMLAVLI
YGIPLRELPIPIGAELISTWACKSRWVVSYLGVFVIVPGSILALTA"
1550..1588
-gene="nptA"
/notes="shared with thyA gene running in opposite
direction"
/function="transcription termination"

```

stem_loop

ORIGIN

Query Match 40.8%; Score 1187; DB 1; Length 1591;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1723 TTAATATAAAGAGTCCCGGAGTGGAGCTTTTATATACAGATAGCTTTAACGC 1782
DB 1591 TTAATATAAAGAGTCCCGGAGTGGAGCTTTTATATACAGATAGCTTTAACGC 1592
QY 1783 TTAAGCGGTTAGGCAAGATGCTGCGGGGATGACGCAAAACACCAATTAAGTACT 1842
DB 1531 TTAAGCGGTTAGGCAAGATGCTGCGGGGATGACGCAAAACACCAATTAAGTACT 1472
QY 1843 CACCAACCAATTTGCTCTTACAAAGCCCAAGTTAGATGAGCTCAGCACCCTTAATAG 1902
DB 1471 CACCAACCAATTTGCTCTTACAAAGCCCAAGTTAGATGAGCTCAGCACCCTTAATAG 1412
QY 1903 CAGTTGGGTTAAGAAAGAAATTCCTTAATCAAGACCGTAGCCTCAAGTTAAAGCTTAA 1962
DB 1411 CAGTTGGGTTAAGAAAGAAATTCCTTAATCAAGACCGTAGCCTCAAGTTAAAGCTTAA 1352
QY 1963 GTGCAACGACGCAATTTGCAAGCAACGCGCAAACTCACAGAGACAGCGGTGGCGC 2022
DB 1351 GTGCAACGACGCAATTTGCAAGCAACGCGCAAACTCACAGAGACAGCGGTGGCGC 1292
QY 2023 GAGCAGACGAGTAATACAGTGCCTAATTTGGCAGCTTAAGTAAATGGGTAGATTTCAG 2082
DB 1291 GAGCAGACGAGTAATACAGTGCCTAATTTGGCAGCTTAAGTAAATGGGTAGATTTCAG 1232
QY 2083 CACTTTGACAGCGCAGACGCCACAGAGAGAAACCATTAAGCTGGTGTGTGCGTGAAGA 2142
DB 1231 CACTTTGACAGCGCAGACGCCACAGAGAGAAACCATTAAGCTGGTGTGTGCGTGAAGA 1172
QY 2143 TTGAATTAATACCGTAACCACTGTACTGAAGCAATACCGTGTAGTGGCCCTCGGCCAAT 2202
DB 1171 TTGAATTAATACCGTAACCACTGTACTGAAGCAATACCGTGTAGTGGCCCTCGGCCAAT 1112
QY 2203 CGCATTTTGTGAATTTTACAGGTGGCGGCCCAACCATCAAACTCTTCACTCACTTTGCCCAT 2262
DB 1111 CGCATTTTGTGAATTTTACAGGTGGCGGCCCAACCATCAAACTCTTCACTCACTTTGCCCAT 1052

```

```

QY 2263 CACCGTAATGCGACGAAATGCTGCAATACCAATACGATAAGTGGACACCAACGAA 2322
DB 1051 CACCGTAATGCGACGAAATGCTGCAATACCAATACGATAAGTGGACACCAACGAA 992
QY 2323 AGTATTACCAATACCGAAAGCTGGTTTCAAGCCCTGTGTGACAGTTTGGTAATCGG 2382
DB 991 AGTATTACCAATACCGAAAGCTGGTTTCAAGCCCTGTGTGACAGTTTGGTAATCGG 932
QY 2383 TTTGATAAATCAAAACCTTTCATGCTCATATCCGACGTCGCAAGACGAGCGAAGAG 2442
DB 931 TTTGATAAATCAAAACCTTTCATGCTCATATCCGACGTCGCAAGACGAGCGAAGAG 872
QY 2443 CCAGTGTGACGCTTCTCTTAATAAGCCAAACATATTTCTTAGAGTGTGAAGATCAGAC 2502
DB 871 CCAGTGTGACGCTTCTCTTAATAAGCCAAACATATTTCTTAGAGTGTGAAGATCAGAC 812
QY 2503 CGCAGAAAGATTGAATAATCGTGTGATGCTGACACTGGCGCAAGCAACGCGAACTCTTC 2562
DB 811 CGCAGAAAGATTGAATAATCGTGTGATGCTGACACTGGCGCAAGCAACGCGAACTCTTC 752
QY 2563 TTTACAGCGCATATGCGCAAGGCTGACGAGAGTATTGTGACAGTAGTACCAATATTGGC 2622
DB 751 TTTACAGCGCATATGCGCAAGGCTGACGAGAGTATTGTGACAGTAGTACCAATATTGGC 692
QY 2623 ACCCATACCATAGAGATCGGGGTTTCAACCGGTAAACCGCGCAACGAGACCAACAT 2682
DB 691 ACCCATACCATAGAGATCGGGGTTTCAACCGGTAAACCGCGCAACGAGACCAACAT 632
QY 2683 AATGAAGTACCGTGCTTGAAGATTTGAATCAGTCCGTTGCCATTAACCAATCATCAA 2742
DB 631 AATGAAGTACCGTGCTTGAAGATTTGAATCAGTCCGTTGCCATTAACCAATCATCAA 572
QY 2743 TCCGCAATTTGGGTTGGGAAACCAATTCATAATCTTTGGCTTGAATCCCGGTTGCCA 2802
DB 571 TCCGCAATTTGGGTTGGGAAACCAATTCATAATCTTTGGCTTGAATCCCGGTTGCCA 512
QY 2803 TTTAAACCGCTGCCGACCATTCGCGACTGCAAGAGTAGTAATATACATGAAGGCAA 2862
DB 511 TTTAAACCGCTGCCGACCATTCGCGACTGCAAGAGTAGTAATATACATGAAGGCAA 452
QY 2863 GTTTGCCCAACGTAGGCTTTGCTGTGTCAGCGCAATTCGCGCTGCAG 2909
DB 451 GTTTGCCCAACGTAGGCTTTGCTGTGTCAGCGCAATTCGCGCTGCAG 405

```

RESULT 8
 AP005074 298900 bp DNA linear BCT 05-MAR-2003
 DEFINITION Vibrio parahaemolyticus DNA, chromosome 1, complete sequence, 2/11.
 ACCESSION AP005074 BA000031
 VERSION AP005074.1 GI:28805287
 KEYWORDS
 SOURCE Vibrio parahaemolyticus
 ORGANISM Vibrio parahaemolyticus
 Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 Vibrionaceae; Vibrio.

REFERENCE
 AUTHORS Nesu,H., Iida,T., Sugahara,T., Yamaichi,Y., Park,K.S., Yokoyama,K.,
 Makino,K., Shinagawa,H. and Honda,T.
 TITLE A filamentous phage associated with recent pandemic Vibrio
 parahaemolyticus O3:k6 strains
 JOURNAL J. Clin. Microbiol. 38 (6), 2156-2161 (2000)
 MEDLINE 20295086
 PUBMED 10834969

REFERENCE
 AUTHORS Makino,K., Oshima,K., Kurokawa,K., Yokoyama,K., Uda,T.,
 Tagomori,K., Iijima,Y., Najima,M., Nakano,M., Yamashita,A.,
 Kubota,Y., Kimura,S., Yasunaga,T., Honda,T., Shinagawa,H.,
 Hattori,M. and Iida,T.
 TITLE Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
 distinct from that of V. cholerae
 JOURNAL Lancet 361 (9359), 743-749 (2003)

MEDLINE
 PUBMED
 22508454
 12620739
 3 (bases 1 to 298900)
 AUTHORS
 Oshima, K., Kurokawa, K., Nakino, K., Yokoyama, K., Yasunaga, T.,
 Honda, T., Shinagawa, H., Hattori, M. and Iida, T.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (09-APR-2002) Ken Kurokawa, Osaka University, Genome
 Information Research Center, 3-1, Yamadaoka, Suita, Osaka 565-0871,
 Japan (E-mail: ken@gen-info.osaka-u.ac.jp)
 URL: <http://www.gen-info.osaka-u.ac.jp/>, Tel: 81-6-6879-8365,
 Fax: 81-6-6879-2047)
 COMMENT
 This clone was isolated from a patient presenting with acute
 gastroenteritis.
 FEATURES
 source
 1. 298900
 /organism="Vibrio parahaemolyticus"
 /mol_type="genomic DNA"
 /strain="O3:K6"
 /sub_strain="RIMD 2210633"
 /db_xref="taxon:670"
 /chromosome="1"
 115..315
 /gene="VP0305"
 115..315
 /gene="VP0305"
 /note="similar to PIR:S56442 percent identity 49 in 63 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="hypothetical protein"
 /protein_id="BAC58568.1"
 /db_xref="GI:28805288"
 /translation="MLREPAVYRPLQVAFVYTLFPGQFFIVGVSFCEDNKKLLPD
 VKQKRLAVREYNQALSLSA"
 complement (399..1037)
 /gene="VP0306"
 complement (339..1037)
 /gene="VP0306"
 /note="similar to GB:AAF95690.1 (AE004324) percent
 identity 74 in 208 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="peptide methionine sulfoxide reductase"
 /protein_id="BAC58569.1"
 /db_xref="GI:28805289"
 /translation="MLNKOTLISIDALPGRBOPMOIEDCHFNOSLTAFLAHQOQ
 ILGMCGFCAERLFWQIDGVSTSVGAGTTPNPTVEVCTGKTHYVRYVDE
 RVLSIAGLAVFWEKDPDQMGQNDICTORSALTYTSODQSLAKSKLQYQAL
 EAEKSTITTEIVPAGPYFAETTHQOYLAKNPDGTGIGTGCFFPSLQ"
 1205..2917
 /gene="VP0307"
 1205..2917
 /gene="VP0307"
 /note="similar to GB:AAF95689.1 (AE004323) percent
 identity 66 in 571 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="BAC58570.1"
 /db_xref="GI:28805290"
 /translation="MIKALPIIVGLLVVSQAVADVSLKLGIDALDNVAVLSS
 IPEKDVSTSRPQALDOSTEALNALGYHAKISYISGNDLIVNIHAGLPKIK
 VMDVYISGEAKDEEFPNLAKSLPYGRLINQGERDSKSTRNLALORGYPNDRK
 LAKLEVLPELNAKAVRLHDSGRIRHGEVETLGSQIMENRVSRRPEIGPFLVSD
 VGEYNQMLSTWDFSSVVEPDISKEDRELPKIVSLAPAKNQLETGIGLSTDTGV
 RGTLMKKPWSARSHSFNALSISKEPQITAGYKIPLDVLRVYQLOGLKMLDN
 RDTESLNSLAVRHWLTDGQWHTVYHLYENFSQIQDDGVQFVLPFGATFSTRY
 RCGSMWMBKQSVTEYVGPALSETRVLRLLGRSSWIRGIGENHRLPREGGANT
 TEPEKLSPELRFAGDNNIRGYESIIPVDEGALGAKYILISTLEYRYVGN
 WMAATFYDIDANDPTPEKMSGAGVIRWASPVGPSDFANGIDENKNNFRHFHS

gene
 CDS
 2914..6681
 /gene="VP0308"
 2914..6681
 /gene="VP0308"
 /note="similar to GB:AAF95688.1 (AE004323) percent
 identity 54 in 1258 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="BAC58571.1"
 /db_xref="GI:28805291"
 /translation="MIKKMFWTKWLSLIGLLLLITIFVATVLFTHPGKALIMGA
 EKALPOLKVENIQGSLFPRFALONVSFDESILHIDAKYERATALINERCFDPKVCYD
 EALQGVNFMPELPASPEPEETPLPSISTPVPFVFNKSVENDIVANLQIDW
 QSTSLASQGRVIAIAPALKDINVAIAPASBAEBAEAEKEDATPOOIVPEVW
 IPIVVARLDIHNKLAGEPVIVNIHGLIVARAGGVYDKTSELMEVEGKLSIQ
 VTSADPYKAKDLALVQADKQKSLSLASGSVGDLSLNATLSLRYOATIKNDIOP
 LKTDLPFVIVNVOQMPFPGSDIOVSFSLAKGSLDGEYALERTKAGKDIPIV
 DVALLNGKTLQIDLESIVTKLGSLSGKVMWMAADINQADINLKNIQPGQWES
 AEGDISGSLSTSGSLTEQGMQVSLPKLIDIGILNGYPLINNGQLASDNKGEIDIO
 LITQGLALSHGPNLSAGKIDKQIMDEVEVPPDPKASVPLAGKQKVALGSLK
 EPDINDILANQINMOQANVTITLKQDVTPLPAPKANLSLANNITVYDIDKVSAD
 LEVSGDEKHLQILDVSDVLSISLEIGTKPKKEMTMDGALRLRLTISQOGPMSIQ
 KSTAVKVTIKQILANVQAHCWLOAKSSVCLTEDISVGTGQAKLAINDPNDOKPL
 PERTKQSSVNAQAFKAPKREKPEVTLSVNNPKQVQALAEQPIKNDPESFKAL
 AKDKDAEWLPVKONGDLSGKVSILNVSSEPTIDNGVASTHLPFLALIEYSL
 FFAANNMTLALSQVWHPKRVNQFLIDMKLQGEVTPIDINSQVINFKHQQDLVA
 GILTPDGKLEITGDADWDRLODMRTTRPAKQELKVMPEVVKIYKEDMTIDTPOL
 AKVEGAINLPKQRIVIBELPSAVGSDDTVILNKDLOPVEVAMPEVETDINIKI
 GDDPOLAFLGKGLKSLNVTOKKQGFIVGEVNIIVGSRSGQDVLIEGKILNN
 GPADQPVYSIKAIRNDMTODVIGAVTAPSPVSEIPEEAMPANLSTLLG
 CDIDSSGQNAVTTLTLLIGLSLAKSGRVGELGEPAGVQDLOLDAGSGDSQVAVSVY
 ILPGLOVKYGVGIRNSLGEFTYRRLMODLILBAVSGVDSAVDILYQEPFP"
 6726..7073
 /gene="VP0309"
 6726..7073
 /gene="VP0309"
 /note="similar to GB:AAF95687.1 (AE004322) percent
 identity 64 in 115 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="BAC58572.1"
 /db_xref="GI:28805292"
 /translation="MOHLIFVYGLTKGEVNHVYLSAQFLGHEHTDAQFSLVYGSY
 PLASGQRSLQGEVYLIDDTLVLDKIEDVPEVYRRSALTPEFGQWYLYQDTQL
 TREIASGDMQGV"
 7170..7532
 /gene="VP0310"
 7170..7532
 /gene="VP0310"
 /note="unknown"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="hypothetical protein"
 /protein_id="BAC58573.1"
 /db_xref="GI:28805293"
 /translation="WRVLIIGAMMLLSACAQTLTPASTVTDQWAGKOSALDGLRE
 LSQERIALKIDGTNHAPELIIAYQNGYQEGQYCEQSAVYLGVVGPYFGICDDVDP
 FPOHYDAQMSACAP"
 complement (7607..8137)
 /gene="VP0311"
 complement (7607..8137)
 /gene="VP0311"
 /note="similar to GB:AAF95686.1 (AE004323) percent
 identity 90 in 176 aa"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1

```
/product="inorganic pyrophosphatase"
/protein_id="BAC58574.1"
/db_xref="GI:28805294"
/translation="MSLNHVPACKSLPEIDYVIVLIPANADPIKEVDKDSGAVEVDR
FKSAPMFPCNTGYNNITLSDSDPVDVLPPIYPLMPSVIRCRPVLKMTDSGE
DAKVAVPSKISKEYEHIQDVGDIPELLKAQITHFPERYKELESKMVKVDMADVE
AAKAEIIQSYERAKQNK"
/complement(8265..3281)
/gene="VP0313"
/complement(8265..3281)
/gene="VP0313"
/complement(8265..3281)
/notice="similar to GB:AAF95685.1 (AE004323) percent
identity 88 in 336 aa"
/codon_start=1
/evidence=not experimental
/translation_table=11
/product="fructose-1,6-bisphosphatase"
/protein_id="BAC58575.1"
/db_xref="GI:28805295"
/translation="MSGMRITGEFIVEKQADFPHPASGDLSLHSLRIAKIVREIN
AAGLDITGAVGTENVOGBAQOKLDVYANDKEKALFADDOYCGVASEDEDAVFNK
ELNOKAKYVLMPLDSDSNIDVNSVGTIFSIYRSPISGIPATEEDPLOGHVOVA
AGVIVGSSITMLVYTTNGNMGPTVDPISGSRCLSHENMTPEDEKTIYSINGNITRF
POGVKVIKYCCOENVPEDRPYTSRISGLVADFRHLNKGIVLPSTOSHPOKLR
LYECNPMAFLIEQMGIIASDGVNRIMDIKPELHORVPFVGSNMVAKVEFLEIH
RDEE"
/gene="VP0312"
/complement(9265..9360)
/gene="VP0312"
/complement(9265..9360)
CDS
35.2%; Score 1052.4; DB 1; Length 298900;
Query Match Best Local Similarity 76.6%; Pred. No. 0;
Matches 1303; Conservative 0; Mismatches 391; Indels 6; Gaps 1;
```

```
QY 602 GGTAAACCTCGTCGCTAGGAGCGATATCCGACTGTTTTAGCTGATGATACGATATC 661
DB 240605 AAAAGCCTCGCCCAATGGTTCGGATACAGGGCTATTTTAGCTGATATGATCATTC 240664
QY 662 CGCTCCCTTGGGATATAGTCGCGAGGCAGATGCTGAGTGGGTCTGTTGGGCTC 721
DB 240665 CGCTTCTTAGTCGAGTTGCTTCGCGAACGATGCTGAGCTGGGTCTATTTGGTGTAT 240724
QY 722 ATTCAATGGGCAAAATCCTCTCTTACTATGATGATCATCGGATTTGATGATGAT 781
DB 240725 ATTTCGATGGGCAAAATCTATCATGATGATGATGATGATGATGATGATGAT 240784
QY 782 TGGTCTTACAGCCCGGTTGTATCAAGAC-----CGTGAAGACAAATAGGCTAGT 835
DB 240785 TGGGCTTACAAACGTCGCTTTTACCAAGACAAAGCTCAAGTTAAACGATTAAGTAT 240844
QY 836 TAGGTAAACAGTATTTAGATCTTTGTACGCGATCGTCATCAAGGTTGGGTTGAA 895
DB 240845 GGTGGAAGAGATTTTATGATTTGTGTCAAGCATCGTTATCAAGGTTGGGTTGAA 240904
QY 896 AATGAAGCAAGCGCAAGCGTTGTTGACTGTATTAATGCCATTTGACCTACGATGTG 955
DB 240905 AATGAAGCAAGCGTAAAGCGTTGCTGACGCTATCAAGCGGATTTAAACCTACGATGT 240964
QY 956 GGCAACATCAGTTCCCTCTAGTACTACAGCAAGAGTTTGGAAAGCGCCGTAGCC 1015
DB 240965 GCGAACATCATTTCCATTCCTATGATGATGATGATGATGATGATGATGATGATGAT 241024
QY 1016 GAGTTGCTCGGCTATTTATGCTGTTAGATTAATGCGCGGATTTTCGCAATTAGTACC 1075
DB 241025 GAGTACTTGGCTATTCATTCGCTGTTATGATTAATGCTGAAGATTTCCGTAATTAGCAC 241084
QY 1076 AAAACCTGGAGTGGTATGCGCAATTTAAACCAAGCATGCTCAACATCTTACCGTAA 1135
DB 241085 AAAACCTGGAGTGGAGCGCAACCTTAACCAAGCATGCTGTTGAATTAACCATACCGTAA 241144
QY 1136 GGTGAGATGATGATGAGAGCGCTGATGATGATGATGATGATGATGATGATGATGAT 1195
DB 241145 GGTGAGATGATGATGAGAGCGCTGATGATGATGATGATGATGATGATGATGATGAT 241204
QY 1196 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1255
DB 241205 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241264
QY 1256 GCAAGTGAATCTTAACTCTAACATCCGGGTGAATTCACATGAGGAGTGTGGCGCCT 1315
DB 241265 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241324
QY 1316 TGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1375
DB 241325 TGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241384
QY 1376 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1435
DB 241385 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241444
QY 1436 GCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1495
DB 241445 GCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241504
QY 1496 GCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1555
DB 241505 GCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241564
QY 1556 TTCCAGCGCTCAGTTCCATATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1615
DB 241565 TTCCAGCGCTCAGTTCCATATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 241624
QY 1616 TGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1675
DB 241625 TGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241684
```



```

/db_xref="GI:27359909"
/translacion="MSGSEFNLIEKYVGRQAKXVDHLAAGDCLVKPSPNVQJALS
TDTLVCGHFLADADPAWVAHKALASITSLAMGATPEWVSFALTMPLADEAMLAPF
CDGFFKADIVFGIQLIGDITKPELSLITVQGFVEGALITRSARKGDYFVTVGLT
GDAGFADVLIDNSMLRTTRPAQAOLERAKHMSMPRLIAGCALVGLASAIIDSGLIAD
LGHILNRISQVSLDVSFKLIPSESLFLTERMALQVALTSGEEYELCTFYPEARG
TIEBALHVGCKVTCIGQIRPHGFPELHDGCTLMWDLSGFPHFA"
complement (6847..7314)
/locus_tag="VV10318"
/locus_id="AA008848.1"
complement (6847..7314)
/locus_tag="VV10318"
/locus_id="AA008848.1"
/transl_table=1
/codon_start=1
/product="Transcription termination factor"
/protein_id="AA008848.1"
/db_xref="GI:27359910"
/translacion="MGASVKEPAPARNARQFALQAIYSWQITKENVATTEAOLFAGEKY
DEEHSRAEPALAAPEITDVAEPFDLTGVLSHTELDASKIRPYVRPMQDDMMELAL
LRLMYEMTRREDREPYKVINVEALIELAKVFAEDSHKFNVGVLDRAAHVAKK"
complement (7314..7784)
/locus_tag="VV10319"
/locus_id="AA008849.1"
complement (7314..7784)
/locus_tag="VV10319"
/locus_id="AA008849.1"
/transl_table=1
/codon_start=1
/product="Riboflavin synthase beta-chain"
/protein_id="AA008849.1"
/db_xref="GI:27359911"
/translacion="MKVIEGGFPAPNAKIAIVISRFNSINESLSGAIDTLKRGQV
QSDNITVRCGAVELPVAORVAKTKDAIVSLGVSIRGCTPHFDYCEENKKGIA
QSVSEFSLPFAVGELVTDIDCAIERAGTKAGNKGALEALSLMEINVLSEIDS"
complement (7961..9070)
/locus_tag="VV10321"
/locus_id="AA008850.1"
complement (7961..9070)
/locus_tag="VV10321"
/locus_id="AA008850.1"
/transl_table=1
/codon_start=1
/product="GMP cyclohydrolase II"
/protein_id="AA008850.1"
/db_xref="GI:27359912"
/translacion="KPISTFOEIIEDIRLKNVILIMDEDEBNEGDILMAEHITPEA
INMAKYGRGLICTIMTKERCQRLGLPVAQONNAQYTNFTVSLEAEGVATISAA
DKRATVOAAVAKKAKADLVOPGHIFPLAQQGVLTAGTGAEGCDLARLAGPEPS
VIEVLIDNDGTNARRPDLVEVFAEBOGLKGTIADLIERNNTEITIRVACEKLPTEF
GEDIIVAPRTIDNOHIFALPKGTMGDDOTPIRVHLOTFPDLIRSDNRMRSEVTDK
AMRIQEGSVYVILGNESPELLIHRVAFPRQCKEAPKAKKQTSIRVVGSGSOT
LADLGHDMRLSSSKKTHALUGFGLNVEYVCE"
complement (9168..9824)
/locus_tag="VV10322"
/locus_id="AA008851.1"
complement (9168..9824)
/locus_tag="VV10322"
/locus_id="AA008851.1"
/transl_table=1
/codon_start=1
/product="Riboflavin synthase alpha chain"
/protein_id="AA008851.1"
/db_xref="GI:27359913"

```

QY	128	TGGTTGGCCAAATCGCGCAGCGGATCGCGGGCAGGTGTGACCGCGTACGCAAGCTCT	187
Db	215497	TGGCTCGCCAAATCGCGCAGCGCAACAAACGGGCGAGTGGCTGGACGGCAAGCAAGTTTC	215438
QY	188	GACTTGTAATTCGCCGCTTTTTCAGGTAGATCGATGGCGAGTGGTATATGTGATC	247
Db	215437	GACTTGCTGTTCGTGATTCCTTAGGCGGTGTGCTGGGTGGTTCGATTGGGTACGTGCTG	215378
QY	248	TTCTACAAATTTGATCTGTTCCTTCTGACCTCTTTATTTATTCAAAGTGTGACTGCG	307
Db	215377	TTCTAAACTTCGATCTCTCTCTGCTGATTCGAATTTACCTTTAAAGTGTGACAGGT	215318
QY	308	GGCATGTCTCTTCACGCGCGCTTATTTGGGTGTGATCACCGGCATGTTGTGATATGCGGT	367
Db	215317	GGCATGTCTTTCCATGCGCGCTGTGTGGGTGTCTTACGCGTATGCTGTGTATCGGGAAG	215258
QY	368	AAAAACAACGCACTCTTTGTGTGTGCGAATTTGTTCCTTTAGTGTGCATTCCGT	427
Db	215257	AAAAATGCCGAACCTTTTTCGCGTGGCGGACTCTGTTCCTCGTTATGTGCATTTCGC	215198
QY	428	TTGGGATGGGACGATACGTACTTTATGAAATAGTAACCTTTGGGAGAGATTAACGAT	487
Db	215197	TTAGGGGTAGGCGCTTAGGCACTTTATGAACGCGACCTTTGGGAGAGATTCGAT	215138
QY	488	GTGCTTGGGCTTTGTATTCCTTAATGTGTGCCACTGCCGCGCATCTTCACAGCTT	547
Db	215137	GTGCTTGGGCGATGTGTTCACCAACGGCGGACACTACACGCGCACTTCGCAATTC	215078
QY	548	TATGAATTCGCTTAGAAGCGGTGTTTGTCTTTATTTCTTAATTTGTTTATTTGTAAA	607
Db	215077	TATGAGATGSAATTAGAAGGAGGTGTGTGTTCTTTATTCCTCACTGGTTATCCGCAAA	215018
QY	608	CCTGCTCCGCTAGACGACGATACCGGACTGTGTTTTCGTGATACGATACATTCGCTTC	667
Db	215017	CCAGACTTTAGATCCGATCTGTATTCCTGTTTATTCCTGTGTGATAGTATTCCTGTTTC	214958
QY	668	CTTGTGAAATCGTCCGTGAGCCAGATGCTGATTTGGTCTGTGTTGTGTGCTTCATTTCA	727
Db	214957	CTGTGTGAATCGTCCGTGAGCCGAGTGCGAACCTGGTCTGTTGTGTGTTCACTTCA	214898
QY	728	ATGGGGCAAAACCTCTCCTTACCTATGGTATTCATCGGTATTTGATATGATGTTGGTCT	787
Db	214897	ATGGGACAGATCTTGTCTCTCCCGATGATTTATTTGTGTATTTGGCTATGTGCTTGGGCT	214838
QY	788	TACAAAGCGGTTGTATCAAGACCTGTGTACAGCAAAATAGAGGTAGTATGATGAACAG	847
Db	214837	TATAAAGTGGCCACTATCAGAACAAAGTACCGTAAGTAAAGGATTTGAAGTGAACA	214778
QY	848	TATTTAATCTTTGTACGCGCATGTGTGATCAAGTGTTTGGGTGAAATGAAACGAACG	907
Db	214777	TATTTGAATGTGTCGCGATGTGTATGAAGGCGACTGGGTGTGAAGAAAGATGAC	214718
QY	908	GGCAAGCGTGTTCAGCTGTGATTTATGCGCATTTGACCTACGATGTGGCAACATACG	967
Db	214717	GGTAAGCGTCCCTTACTGTGATCAACGCGAATTTAACTACGATGTGCCAACATACG	214658
QY	968	TTTCTCTAGTGACTACACGCAAGATTTTGGAAAGCTGCCGTAACCGAGTTGCTCGGC	1027
Db	214657	TTTCTCTTGTGTACACGCGTAAGATTTTCGAAGGCGGCGTGTGCAGATGTGCTGTGC	214598
QY	1028	TATATTCGTGGTTCAGATAATAGCGCGGATTTTGGCCAAATTAAGTACCAAAACCTGGAT	1087
Db	214597	TACATTCGAGTTTATGACAAATGCGGAAGATTTTGCAGATTTAGTATGAAAACTCGGAT	214538
QY	1088	GCTATGCCAATTTAAACAAGCATGCTCAACAATCTTACCGTAAAGGTAGAGATAC	1147
Db	214537	GCGAAACCCATCTCAATGATGCTTGTGTGAATACCATATTCGCAAAAGCGAAGATAC	214478
QY	1148	ATGGAGCGCGTATGTGTTCAGGATAGACTTGGCTAAGCCTGATGGGTGCATATTT	1207
Db	214477	ATGGGCGGTGCTTAGCGCGTACAAAGTTCGACTTGGGCGAAGCCTGATGGGTGCATATTC	214418

QY 1208 GACCACTGAAAAAGATTGTGATGATTTGAAGCCGTGCGCTTGATGACCGAGGTGAATT 1267
 Db 214417 GACCAATTTGCGTAAATTTGTCGATGACCTGACCGCGGAGATGACATGACCGGGGAGATC 214358
 QY 1268 CTTAATCTTCAAAATCCGGGATGATTTCAATGAGGGGATTTGGCCCTTGATGATGACAC 1327
 Db 214357 CTCATTTCTTCAAAATCCGAGGATTTCAATGAGGGGATTTGGCCCTTGATGATGACAT 214298
 QY 1328 CATCATTTTTCATTTGCTGGGGGATACCTTGATGTCACAGTACTGACGGCTTCAATGAT 1387
 Db 214297 CATCATTTTTCATTTGCTGGGGGATACCTTGATGTCACAGTACTGACGGCTTCAATGAT 214238
 QY 1388 GTGCCCTTGGGATGATTTCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1447
 Db 214237 GTGCCACTGGGCTGATTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 214178
 QY 1448 CAGATCAGAGGAAAAAGCCGGGCTTGCTGATGACAGATGCTGATGATGATGATGATGAT 1507
 Db 214177 CAGATTCAGGAGAAAAAGCCGGGCTTGCTGATGACAGATGCTGATGATGATGATGAT 214118
 QY 1508 CAGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1567
 Db 214117 GAGATTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 214058
 QY 1568 CAGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1627
 Db 214057 ACGTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 213998
 QY 1628 GATGATTTTTCAGTACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1687
 Db 213997 GATGATTTTTCAGTACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 213938
 QY 1688 TAA 1690
 Db 213937 TAA 213935

RESULT 10
 AP005332 249150 bp DNA linear ECT 04-DEC-2003
 LOCUS Vibrio vulnificus YJ016 DNA, chromosome I, complete genome, section 3/14.
 DEFINITION AP005332 BA000037
 ACCESSION AP005332.1 GI:37197435
 VERSION AP005332.1 GI:37197435
 KEYWORDS
 SOURCE
 ORGANISM
 Vibrio vulnificus YJ016
 Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.
 REFERENCE
 1
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 This sequence was determined by the Sequencing Core of the National Yang-Ming University Genome Research Center (YMGRC),
 http://genome.ym.edu.tw/
 Location/Qualifiers
 1..249150
 /organism="Vibrio vulnificus YJ016"
 /mol_type="Genomic DNA"
 /strain="YJ016"
 /db_xref="taxon:196600"
 /chromosome="I"

gene 299..979
 /gene="VV0511"
 CDS 299..979
 /gene="VV0511"
 /note="identified by Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC93275.1"
 /db_xref="GI:37197436"
 /translation="MNVSRKKNKRLTTRKTFGLRPIQTEHGPLYLDYERKENTID
 IADVEPLMAIRVDLRFPEKLKMGSGNVMTDFLRISQSIDHSGKKEGSRVH
 PCVRLJAVRERSSINDHYHLMFNDRFVNLGKTQOENLSSFLVEAMAVAGY
 DYEAGLVHPSKRXNSDSVTHRLNCSDAFDPAFDFPFKHISYLAKENTKHGSGR
 RNRGSHK"
 865..888
 /gene="rpt112"
 865..888
 /gene="rpt112"
 /note="2 units of 12 bp in-tandem repeat"
 /rpt_type=tandem
 /rpt_unit="TGATGCGTTGA"
 1040..1708
 /gene="VV0512"
 1040..1708
 /gene="VV0512"
 /note="identified by Genemark and Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC93276.1"
 /db_xref="GI:37197437"
 /translation="MTHQNLNVAHQRAHTTHDRFNGKRYFYEDDGLVLEYLEIDY
 LTELNKTPRIYAVHINLPSDRGDDYLVFAPFRLISETEBLCKDRTDYTYE
 QQTVIRYITWKASBSTSDHHLILFDRDLFQNGGTGEGKGRFLYKRVKAMSVVE
 AYRQPCSGLGYSPTNEGSELSDAYPEPLQINFEFFRISRLAPITNLDRKXSG
 CSHD"
 complement(1286..1441)
 /gene="VV0513"
 complement(1286..1441)
 /gene="VV0513"
 /note="identified by Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAC93277.1"
 /db_xref="GI:37197438"
 /translation="MKELPIKQDQMYIAGGLRLRPDIANNRLVYSSISGVFLT
 ELISLRF"
 1733..2368
 /gene="VV0514"
 1733..2368
 /gene="VV0514"
 /note="identified by Genemark and Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC93278.1"
 /db_xref="GI:37197439"
 /translation="MPMFTFKKLLMNRVLPSTLDRTFEDYPPYSDKGLYSA
 LMSVDVFOALSCFDLALATRLDLPEDHQPDLISLINDYFNLKELDLDSLFY
 WRKRDVPSHMYRMLPSDYSOHGPAICWDKXNELVGHILKTMOAIAIKHSGEAR
 SLVPLNDGNGVGRKNSKVNINKYTFHMSLSLAMEKRKCFSSSL"
 2436..2642
 /gene="VV0515"
 2436..2642
 /gene="VV0515"
 /note="identified by Genemark and Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="predicted transcriptional regulator"
 /protein_id="BAC93279.1"
 /db_xref="GI:37197440"

QY 728 ATGGGCAAAATCCTCTCTTCTTACCTATGCTGATCATCGGATTTTGTGATGATGCTTGGCT 787
Db 175867 ATGGGACAGATCTTCTCTCTTCTTACCTATGCTGATCATCGGATTTTGTGATGATGCTTGGCT 175926
QY 788 TACAGCGCGGTTTGTATCAGACCGGTGTAGCAGCAAAATGAGGTAGTGGTGAACAG 847
Db 175927 TATTAACCGTGGCCTATCAGACCAAAATGAGGTAGTGGTGAACAG 175986
QY 848 TATTAGATCTTTGTCTGACGCGATCGTCGATCAAGGTGTTGGGTGAAATGAAGCAAGC 907
Db 175987 TATTAGATCTTTGTCTGACGCGATCGTCGATCAAGGTGTTGGGTGAAATGAAGCAAGC 176046
QY 908 GCGAAGCGTGTGTTTACTGTGTATTAATGCGCATTTGACCTGAGTGGGCAACATCAG 967
Db 176047 GGTAGCGCGTGTGTTTACTGTGTATTAATGCGCATTTGACCTGAGTGGGCAACATCAG 176106
QY 968 TTTCCTAGTACTACACGCAAGATTTTGAAGGTGCGGTAGCGAGTTCCTGCGC 1027
Db 176107 TTTCCTAGTACTACACGCAAGATTTTGAAGGTGCGGTAGCGAGTTCCTGCGC 176166
QY 1028 TATATTCGTGTTACGATTAATGCGCGGATTTTGGCCATTAAGTACCAAACTGGGAT 1087
Db 176167 TATATTCGTGTTACGATTAATGCGCGGATTTTGGCCATTAAGTACCAAACTGGGAT 176226
QY 1088 GGTATGCGCAATTTTAAACCAAGCATGCTCAACATCTTACCGTAAAGTGAAGATGAC 1147
Db 176227 GGTATGCGCAATTTTAAACCAAGCATGCTCAACATCTTACCGTAAAGTGAAGATGAC 176286
QY 1148 ATGGAGCGCGTGTATGCTTCAAGGTAAGCGTGGGCTTAAGCTGATGCTGATCATTT 1207
Db 176287 ATGGAGCGCGTGTATGCTTCAAGGTAAGCGTGGGCTTAAGCTGATGCTGATCATTT 176346
QY 1208 GACCGATGAAAAAATGTTGATGATTTGAGCGGTGGCGTGTATGACCGAGTGAATTT 1267
Db 176347 GACCGATGAAAAAATGTTGATGATTTGAGCGGTGGCGTGTATGACCGAGTGAATTT 176406
QY 1268 CTTAATCTTACATCCGGGTGAATTTCAATGAGGCTTTGCGCTTGCATGATGACG 1327
Db 176407 CTTAATCTTACATCCGGGTGAATTTCAATGAGGCTTTGCGCTTGCATGATGACG 176466
QY 1328 CATCATTTTTCATTTGCTGGGGGATACCTGTTATCTCAACAGTACACGCTTCAATGAT 1387
Db 176467 CATCATTTTTCATTTGCTGGGGGATACCTGTTATCTCAACAGTACACGCTTCAATGAT 176526
QY 1388 GTGCCCTTGGGCTGTAATTTCAATGAGTGTGAGTGTATGCTTCTTGGCGTGAAGCA 1447
Db 176527 GTGCCCTTGGGCTGTAATTTCAATGAGTGTGAGTGTATGCTTCTTGGCGTGAAGCA 176586
QY 1448 CAGATCAAGGAAAAAGCGGCTTGGGCTTCAACAGTGTGATGATGCTGATGATGAT 1507
Db 176587 CAGATCAAGGAAAAAGCGGCTTGGGCTTCAACAGTGTGATGATGATGATGATGAT 176646
QY 1508 CAAGATCACTGAAATGATGCGCGATGCTGAGTAAACGTCAGCCATTCGACGCGCT 1567
Db 176647 CAAGATCACTGAAATGATGCGCGATGCTGAGTAAACGTCAGCCATTCGACGCGCT 176706
QY 1568 CAGTTCATATCAATCAAAAGATTAAACATGCGAGATTTGAAACTTGGCTCATTTG 1627
Db 176707 CAGTTCATATCAATCAAAAGATTAAACATGCGAGATTTGAAACTTGGCTCATTTG 176766
QY 1628 GAGATTTTGAAGTCAAGCGATATCAAGTTCAGATTCATTCATCAACCGCTTCACTG 1687
Db 176767 GAGATTTTGAAGTCAAGCGATATCAAGTTCAGATTCATTCATCAACCGCTTCACTG 176826
QY 1688 TAA 1690
Db 176827 TAA 176829

RESULT 11
AX020429 852 bp DNA linear PAT 07-SEP-2000
LOCUS
DEFINITION Sequence 1 from Patent WO935271.

ACCESSION AX020429
VERSION AX020429.1 GI:10044145
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..852
/organism="Vibrio cholerae"
/mol_type="unassigned DNA"
/db_xref="taxon:666"
1..852
/note="unassigned protein product; thylA nucleotide sequence"
CDS
1..852
/codon_start=1
/transl_table=11
/protein_id="CAC07770.1"
/db_xref="GI:10044145"
/db_xref="EMBL:CAC07770"
/translation="PROYLDCRILVDCGWNENRTGKCLTVINADILYDVGNCF
PLVTRKSPKAAVAELIGIRGYNDAPFRLGRTDANALNQLMNPYRGED
DMGRYGVGRMAKPDGSHIDQLKRIYDLSRGVDBRGLNLFNPGEPHGLRPLC
MYSHPSLIGDYLINLSTORSCDVLGNLNPVQYVPLAMAOITGKPKGLAHRKIV
NRIYQDLELRNDVQLKREPPAPQFHPKIKLIDLEWTVLDDPDVICYOHP
IYPSV"
misc_difference
1..3
/note="Xaa= Methionine"
ORIGIN
Query Match 28.9%; Score 841.4; DB 6; Length 852;
Best Local Similarity 99.3%; Pred. No. 1.7e-249;
Matches 845; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 840 TGAACAGATTTAGATCTTTGTCAGCGCATGCTGATCAAGGTGTTGGGTGAAGAATG 899
Db 2 TGAACAGATTTAGATCTTTGTCAGCGCATGCTGATCAAGGTGTTGGGTGAAGAATG 61
QY 900 AACGAACGGGCAAGCGTGTGTTGATCTGATTAATCCGATTTGACCTTACATGTCGCA 959
Db 62 AACGAACGGGCAAGCGTGTGTTGATCTGATTAATCCGATTTGACCTTACATGTCGCA 121
QY 960 ACATCAGTTTCTCTAGTACCTACACGCAAGGTTTGAAGCTGCGGTAGCCGAGT 1019
Db 122 ACATCAGTTTCTCTAGTACCTACACGCAAGGTTTGAAGCTGCGGTAGCCGAGT 181
QY 1020 TGCTCGGCTATTTGCTGTTTACGATTAATGCGCGGATTTTGGCCATTAAGTCAAAA 1079
Db 182 TGCTCGGCTATTTGCTGTTTACGATTAATGCGCGGATTTTGGCCATTAAGTCAAAA 241
QY 1080 CTGGGATGCTTAATGCAATTTAAACAGACAGTGTCAACATCTTACCGTAAAGG 1139
Db 242 CTGGGATGCTTAATGCAATTTAAACAGACAGTGTCAACATCTTACCGTAAAGG 301
QY 1140 AGATGACATGGAACGCGTGTATGCTTCAAGGTTAGAGCTTGGCTAAAGCTGATG 1199
Db 302 AGATGACATGGAACGCGTGTATGCTTCAAGGTTAGAGCTTGGCTAAAGCTGATG 361
QY 1200 GTCATATTGACAGTTGAAAAAATGTTGATGATTTGAGCGCTGGCGTTGATGACGAG 1259
Db 362 GTCATATTGACAGTTGAAAAAATGTTGATGATTTGAGCGCTGGCGTTGATGACGAG 421
QY 1260 GTGAATTTCTTAATCTTCAATCCGCGGTAATTTGACATGAGGCTTGGCGCTTGA 1319
Db 422 GTGAATTTCTTAATCTTCAATCCGCGGTAATTTGACATGAGGCTTGGCGCTTGA 481

QY 1320 TGTACAGCATATTTTTCATTTGCTGGGGGATACCTTGATCTCAACAGTACTACGCGTT 1379
DB 482 TGTACAGCATATTTTTCATTTGCTGGGGGATACCTTGATCTCAACAGTACTACGCGTT 541
QY 1380 CATGTGATGTGCGCTTGGGGGTTGAATTTTCAACATGTGAGGTTTATGTTTCTTGGGC 1439
DB 542 CATGTGATGTGCGCTTGGGGGTTGAATTTTCAACATGTGAGGTTTATGTTTCTTGGGC 601
QY 1440 TGTATGGACAGATTCACAGGGAAGAAAGCCGGGCTTGGCGATACAGATTCGTCATTCGCC 1499
DB 602 TGTATGGACAGATTCACAGGGAAGAAAGCCGGGCTTGGCGATACAGATTCGTCATTCGCC 661
QY 1500 ACATTACCAAGATCACTCGAATTGATGCGCGATGTGAGCTAAAGGTGAGCCATTC 1559
DB 662 ACATTACCAAGATCACTCGAATTGATGCGCGATGTGAGCTAAAGGTGAGCCATTC 721
QY 1560 CAGCGCTCAGTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1619
DB 722 CAGCGCTCAGTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 781
QY 1620 TCACTTGGATGATTTTGAAGTCAAGGATACGATTCAGTTCAGATTCATTCATTCATTC 1679
DB 782 TCACTTGGATGATTTTGAAGTCAAGGATACGATTCAGTTCAGATTCATTCATTCATTC 841
QY 1680 TTTCACTTAA 1690
DB 842 TTTCACTTAA 852

RESULT 12
LOCUS AX009482 838 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 2 from Patent WO961634.
ACCESSION AX009482
VERSION AX009482.1 GI:9996767
KEYWORDS
SOURCE Vibrio cholerae
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.

REFERENCE
AUTHORS Carlin, N. and Lebens, M.R.
TITLE Method of producing thy a<->strains of vibrio cholerae, such strains and their use
JOURNAL Patent: WO 961634-A 2 02-DEC-1999;
CARLIN NILS (SE); SBL VACCIN AB (SE); LEBENS MICHAEL R (SE)
FEATURES
source 1. 838
/organism="Vibrio cholerae"
/mol_type="unassigned DNA"
/db_xref="taxon:666"

ORIGIN
Query Match 28.8%; Score 838; DB 6; Length 838;
Best Local Similarity 100.0%; Pred. No. 1.9e-248;
Matches 838; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 TGTATCTTCTACAAATTTTGTATCTGTTCTTGTGCTGACCTTTTATTTATTCAAAGTGTG 300
DB 241 TGTATCTTCTACAAATTTTGTATCTGTTCTTGTGCTGACCTTTTATTTATTCAAAGTGTG 300
QY 301 GACTGCGGAGTGTCTTTCACAGCGCGCTTATTTGGGTTGATATACCGCATGTTCTGGTA 360
DB 301 GACTGCGGAGTGTCTTTCACAGCGCGCTTATTTGGGTTGATATACCGCATGTTCTGGTA 360
QY 361 TGGCGTAAACCAACAGCACCTTCTTTGATGTGGCCGATTTTGTGCTTATGTC 420
DB 361 TGGCGTAAACCAACAGCACCTTCTTTGATGTGGCCGATTTTGTGCTTATGTC 420
QY 421 ATTGCGTTTGGGATGAGGATTCGGTATCTTATGAAATGTAATGTAATTTGGGAGAGT 480
DB 421 ATTGCGTTTGGGATGAGGATTCGGTATCTTATGAAATGTAATGTAATTTGGGAGAGT 480
QY 481 AAGGATGTGCTTGGGCTTTTGTATTCCTTAATGTGCGCATGCGCGCATCTTC 540
DB 481 AAGGATGTGCTTGGGCTTTTGTATTCCTTAATGTGCGCATGCGCGCATCTTC 540
QY 541 AAGGATGTGCTTGGGCTTTTGTATTCCTTAATGTGCGCATGCGCGCATCTTC 600
DB 541 AAGGATGTGCTTGGGCTTTTGTATTCCTTAATGTGCGCATGCGCGCATCTTC 600
QY 601 TGTAAACCTCGTCCGCTAGGAGCGATTCGAGCTGTTTGTAGTGAATACGGTACAT 660
DB 601 TGTAAACCTCGTCCGCTAGGAGCGATTCGAGCTGTTTGTAGTGAATACGGTACAT 660
QY 661 CCGCTTCTTGTGAATACGTCCTGAGCGCATGCTCAGTTGGGTTGTGGGCTT 720
DB 661 CCGCTTCTTGTGAATACGTCCTGAGCGCATGCTCAGTTGGGTTGTGGGCTT 720
QY 721 CATTTAATGGGCAATCTCTCTTACTTATGATGATCATGATTTTGAATTTGAAT 780
DB 721 CATTTAATGGGCAATCTCTCTTACTTATGATGATCATGATTTTGAATTTGAAT 780
QY 781 TGTGCTTACAGCGCGGTTTGTATCAGACCGTGTAGCAGCAAAATAGGATTTAG 838
DB 781 TGTGCTTACAGCGCGGTTTGTATCAGACCGTGTAGCAGCAAAATAGGATTTAG 838

RESULT 13
LOCUS AE006043/c 10719 bp DNA linear BCT 08-MAR-2001
DEFINITION Pasteurella multocida PM70 section 10 of 204 of the complete genome.
ACCESSION AE006043
VERSION AE006043.1 GI:12720282
KEYWORDS
SOURCE Pasteurella multocida
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Pasteurella.

REFERENCE
AUTHORS May, B.U., Zhang, Q., Li, L.L., Paustian, M.L., Whitlam, T.S. and Kapur, V.
TITLE Complete genomic sequence of Pasteurella multocida, PM70
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
PUBMED 11248100
REFERENCE
AUTHORS Zhang, Q. and Kapur, V.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2000) Department of Veterinary Pathobiology, University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN 55108, USA
FEATURES
source 1. 10719
/organism="Pasteurella multocida"
/mol_type="genomic DNA"
/strain="PM70"
/db_xref="taxon:747"

gene
225..1076
/gene="PM0074"
CDS
225..1076
/gene="PM0074"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02158.1"
/db_xref="GI:12720283"
/translation="MKSEESTCTCVSPPTKAAQIGEDVGVYKATKQIYSPFAIPAG
AFIAIPFVATQONGDAPWGLTKLVGIVFSLGVIMVVCSELFTSTWTAVAR
VSGRIPTROMIKMLVYCGNFGVALFLVLTWPAQIMANQOMGLTIKTOKXIH
HTWDEALGICINIMVCIATWMTYAKTILLDKAVIMILPISMFVAGSPHSANM
IPMGMLAHFASPEFMQAVSLNPAQFDLMDTHFVYKNLLPVTIGNVGNALCVALTQ
WYINRPH"
gene
1156..3486
/gene="pf1b"
/note="synonym: PM0075"
CDS
1156..3486
/gene="pf1b"
/codon_start=1
/transl_table=11
/product="pf1b"
/protein_id="AAK02159.1"
/db_xref="GI:12720284"
/translation="MSMSQLMETQCKAWESPTGGDMOTENVNRPDIQKNYTPYEGDES
FLADTERTKLMDVMKIKYENKTKHPYIDDDTSTITSHAGYIDSLKXIVGL
QDAPLKRAMPFGGINVSKSCVYRRELPEYBQIFTEYKTHNGCVDPYVPTDIL
RCRKGVTIGLPAAYGRRIIGDYRPMALYGADFLMKENQPTSLQDKLREGSDIA
TIOLEEEIAEOHIALGKMEKMAASYDISEPATMEAVQWTFAYLAIVKSNQGA
MSFGVSTFLDIYERDLKAKITEOEAOELIDHLWMLRMVRELRTPEYDQFLSPGB
WMAETTLAGMGLDRLTVTKNSFRILHTLYMGSPSEBNITLISEKLPEGFKYAAK
VSIPTSSVOYENDLGRMPDFONDYALACVSPMVGKMOFGFARMLAKTLLYAN
GVDEKSGDQVQPKTDPTITSEYLDVDPMTSELMDLAKOYATLNIHFMMDKXA
YEALMALHDRVFRITACGTAGUSVADSLSAIKYAKYVRGDIETLKXKAGVQVI
AKDVAIDELIEGEYPOFGNNDNRVDEIACDIVERPKTIQKLGITRYNATPTOSLTIT
SNVYVQKTKGTPDGRGAPFGANPMHGRDKGAVASLTSVAKLPYAKUGISYI
TSPVFNALGKDYEAQKRNLAGMDGYFHHEATIEGGQHLNVMNRMELDMANENE
KYPOLITRVSQVAVFNSLTKEQQDQVITRTFTQAM"
gene
3659..5698
/gene="est"
/note="synonym: PM0076"
CDS
3659..5698
/gene="est"
/codon_start=1
/transl_table=11
/product="Bst"
/protein_id="AAK02160.1"
/db_xref="GI:12720285"
/translation="MKINTLSFALFCSLSAALANDVVFQDSJSDMGQCTWNRKAST
LRADGSYHLVDEYLAQAFGKLIPTQGLNVAISGVIVGANTRTAEQPHIALEX
QINEYLAHVKKALHILWAGNDLATVATVYKTPTEKQAVYIASINMAQTMQ
OMGALQOAGVNOIIAPTIENVYTPPEFDKGEAAGIOAKSYGLIKOSDFVAFKA
AAKQILQPTKRLIEPEKRRVOTLEKAAEDFNSRYFTQIALSAGVNAIGIETLI
KEKTIYEQAOATAFNAITSITSALNQVQGNIVRTDIDGLKDMMTTAEYGLNNTT
VCKESTADPACPKPEPDQNASORLEPDSVHPDITAHKAKSYTILANLOTPTDMQ
LQIQVOQOTELADPTRTESNHRLOQOQSAQSVDTLAVYQKQSGSHVAKQVQFN
QMOIALVAAQKQODVQGLININTONRVYSLIRYDANWMLGVLQVNNVTITTDYR
AYIGHAHOQSAETELESFSGMLPAGYEWEKFKALGIALIADIKTKTQVNAFTGBNRG
ITQMBQARTKESVSGVGFDRYQAAQWQPVYTRVWVEMNDIPTIOALINSGSEM
TVLPDQDNQWNI VAGLHKPISSGFHNLAIRDLGRKNMLSNITTLQAGIGTFE"
gene
5862..6602
/gene="act"
/note="synonym: PM0077"
CDS
5862..6602
/gene="act"
/codon_start=1
/transl_table=11
/product="Act"
/protein_id="AAK02161.1"
/db_xref="GI:12720286"
/translation="MAVVGRIHSYSQGTVDGPIRIFILFMOGLMRCQYCHNRDITW
LHAGKEITVEILMKEVIVVYRHHFNASGGGVATSGGAVALQAEFVADMVSACKKEGIIHT

gene
CLDITNGFVRNVDHVIDELDVTDLVDLIDKQNDKRIHQNLIGENKRTLEBPAQYLAAR
NOPWIRVYVVGTYDDEDDHILGHFIKDKKNIEKVELLYPHRLGAHKWEMGEKYE
LEEVNPTKSEIENHKSLSESGHIIK"
gene
complement (6819..7319)
/gene="PM0078"
complement (6819..7319)
/gene="PM0078"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02162.1"
/db_xref="GI:12720287"
/translation="MQSDMDERKMRHLRLADKAEQGEIVGAVLVDEQGNLGEW
NLISIEDPTAAIEIVLROGQORLONRYLNTLYTLTLPQTCQAGHILSRIRLV
IPASDYKTGAQVSRPHFPDDYKMHNAIEITAGVLOECSQKISAFQKRRBOQKXIAQ
KSHLFT"
gene
complement (7334..8185)
/gene="chya"
/note="synonym: PM0079"
complement (7334..8185)
/gene="chya"
/codon_start=1
/transl_table=11
/product="chya"
/protein_id="AAK02163.1"
/db_xref="GI:12720288"
/translation="MKHYIELQRIIVDEGVVENARKRCCLTVIADLTVDVANNQ
PLITRKSXYKAAIAEFLIGYRYGNADFRGLGTYKTDANANENSALNPHRGTD
DMGRVYGVQGRAMRKXENGETVDQIKRIYNNIANGIDDEGEINTFENPEFELGCLRP
KHHPFSLIGDLTYLYSYORSQDVLGNFNOIYVETPLANAOQTGKAGAKYKRII
NNAIYEDQDLMKQNVQLRDPPELQLHINPEIKTLEDELTWVTDKRVGYQCHDA
IKRPSV"
gene
complement (8182..8994)
/gene="lgt"
/note="synonym: PM0080"
complement (8182..8994)
/gene="lgt"
/codon_start=1
/transl_table=11
/product="lgt"
/protein_id="AAK02164.1"
/db_xref="GI:12720289"
/translation="MQSNYFSPPOFDPVFEFBIPIGLRWYGLMILGLFARMLAVK
ANOTSGMWTDDVDSLTFNGEMVGLGRIGVLPYQPDYFLQDPAVYLFVRWGGMSF
HGGILVITLAMILATKIQORQWQADPAFAPILPGLGWRGTFINDELMTGRVDP
WAVLPSPGGILPHRBSQLYEAVYEGIVLPILNWTIKKRPICGATAGFLDGYGFRF
YVFFREPDQGLYFGQHSWGLSTPMILIGAVIMLVAYQSAKREIL"
gene
complement (9004..9798)
/gene="PM0081"
complement (9004..9798)
/gene="PM0081"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02165.1"
/db_xref="GI:12720290"
/translation="MFTFLICIAVAFAGFLAGLFGIGGLVIVPVLYVLMKRYGV
PBMVACALGTSFATIVTFLGSSAQRHKLGNVWDVYKAVIYMTTFISGFEIS
YHKDLSARIKALVIVYLAIKMTLSKRPKPTKALTKASTIAGLIGASSAAGIG
GGIVTFPLNSRGRDKMKAIGSSHACOSLAIKAMSPMIAKRWGGLPEYSGYVILP
ALVGIITLSIFTSIGATVYSLIPVILKKAFLMLILAINMILK"
gene
complement (9801..10397)
/gene="PM0082"
complement (9801..10397)
/gene="PM0082"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02166.1"
/db_xref="GI:12720291"
/translation="MTDPDGYRPNVGIIVCNSKQVYLAARYQGNQWQFPOGGINDNE
SAGQVARELFEWVGLSKMIRYKILPKRLRYDSKVPICIGQKQRMFL
QVADSEKNIWQSSKSPEDGRWVFWIPYQVYSFKKEVYKAKKEFPASVLPDQAK

ORIGIN

ENLISKSNSDLKTHHTTKSTFLTKSHKHFHRSRG"

Query Match 28.0%; Score 814.8; DB 1; Length 10719;
 Best Local Similarity 68.7%; Pred. No. 5,2e-241;
 Matches 1161; Conservative 0; Mismatches 507; Indels 22; Gaps 2;

4 AAGTTTGTATGCTCAGGAGTTATCTGAGTTCCCAATATGACCCCGATGTTTC 63
 Db AAGGATTTTATGCAATGCAATTTTTCATTTCCACAATTTGATCCCGATTTTTGA 8945

64 GATCGCCCTCAGCGGCGCTGGTATGCTTATGATTTGTTGGGTTTCTTTTGC 123
 Db AATCGCTCTATCGGTTTACGTTGATGCTTAATGATCTGCTGGGGTTCTTTTGC 8885

124 TATGTGTGSCCAATCGCCGAGCGGATGCGCGGAGTGGTGAAGCGGTGACAGT 183
 Db CCGTGGCTTGGCGGTAACGCTGTAATCAACAGGCACTGGCTGAACAGACCAAGT 8825

184 CTCTGACTTTTATTGCGCGCTTTTATAGTGTATGATCGGTGGCCGAGTTGATGT 243
 Db GATGAGCCGTTATTTAAACGTTTATGAGCGTTTATTAGGTGACGATTTGTTAGCT 8765

244 GATCTCTACATTTTGAATCTGTCTGCTGACCGCTCTTATTTATTTCAAGTGTGAC 303
 Db GTTGTTTTATCATTCGATTTATTTCTTACAGATCTCGGATTTATTCGTTTGGGA 8705

304 TGCGCGCATGTCCTTCAACGCGGCTTATGAGTGTATGACCGCATGTTCTGGTATGC 363
 Db AGCGGATGTCCTTCCAGTGTGGCTTAATGTTGATTTAGCATTTATATCACAGC 8645

364 GCGTAAAAACAAGCACTTCTTGTGTGTGCGCATTTTGTGCCCCCTTATAGCCATT 423
 Db TAAATATCAAAAACGCGGTTTCTGGCAACGCGGATTTTGTGGCGCGTTATTCATT 8585

424 CGGTTGGGATGAGACGTAATGCTTATGATATGTAATTTGAGGACAGAGTATAC 483
 Db TGCGTTAGTATGGAACGATGCGACATTTTATCATATGACGATTTGTGGGCGGTATAC 8525

484 GATGTGCTGGGCTTTTGTATTCCTTAATGATGAGGCGCACTGCCGCGCATCTTCA 543
 Db GGATGTGCGCATGGGAGTCTTATTCCTCATAGCGGATATTTTACACCCCATTCATCA 8465

544 GCTTATGATTTGCGCTTAGAAGGCGTGTCTGTCTTATTTCTTATTTGTTATG 603
 Db GCTTATGCAAGCGGATTTAGAGGCAATGTGCTATTTCTTATTTGAAATGATATATCA 8405

604 TAAACCTGTGCGTAGGCAAGGATACGGAGCTGTTTATGCTGATACGGTATATTCG 663
 Db AAGCCTGACCAATCGGCGCAACGCTGTTATTTCTTATCTCGGTACGGATTTTTCG 8345

664 CTTCCTTGTGAAATACGCCGAGCAAGTCTAGTTGGTCT---GTTTGTGCGCTT 720
 Db TTTTACGTTGTTTTCGAGCCTGATGCGCAATTAAGGCTTATATTTTGTGCAACA 8285

721 CATTTCAATGGGCAATCTCTCTTACTATGCTATGATATCAATCGGTATTTGATGAT 780
 Db TATCTCCATGGGCAAAATCTCTCCACACCAATGATTTCTCATTTGCTGGGATATATCTT 8225

781 TTGTCTTACAAAGCGCGGTTTGTATCAAGACCGTGTAGACGAAATAGGAGTATAGT 840
 Db AGTGCATATCAAAATG-----CGGTTAAAAAAGGCAAAATATAT 8184

841 GAAACAGTATTTAGATCTTTGTCAGGCACTGTCATCAAGTGTGTTGGTTGAAAAATGA 900
 Db GAAACATATTTAGATCTTTGTCAGGCACTGTCATCAAGTGTGTTGGTTGAAAAATGC 8124

901 AGCAACGGGCAAGCGTTGTTACTGTGATTAATGCCATTTGACCTACGATGTGGGCA 960
 Db AGGAGCTGGCAACGTTGTCTGACCGTGTATGATCGGATCTCACCTACGATGTGGGCA 8064

961 CAATCAGTTTCTCTAGTACTACAGGAGGTTTGTGAAAGTGTGGCGATACCGCATTT 1020

Db 8063 TAAACATTTCCGTTGATCAACACAGCAAAAGCTATTGAAAACGGCAATTCAGATTT 8004

Qy 1021 GCTCGCTATATGCTGATGATAGATATATGCGGAGATTTTGGCAATTAGGTACAAAAC 1080

Db 8003 TTTTGTATATCCGTGCTATGATGCAATGCGGAGATTTCCGAACTCGGACCAAAAC 7544

Qy 1081 CTGGGATGCTATATGCAATTTTAAACCAAGCATGCTCAACAATCTTACCGTAAAGTGA 1140

Db 7943 CTGGGATGCAATATGCTATGATAAAACAGTCATGCTCAATATCCGACCGCAAAAGCAC 7884

Qy 1141 GATGATATGGAAGCGCTGATGTTGTCAGGTTAGCTTTGGCTTAAGCTATAGTGTG 1200

Db 7883 CGATGATATGGAAGCGCTGATGTTGTCAGGTTAGCTTTGGCTTAAGCTATAGTGTG 7824

Qy 1201 TCATATTTGACCACTTGA AAAAGATTTGTTGATTTGAGCGCTGAGGCTTATGATGACGAG 1260

Db 7823 AACCGTATCATATTTGCGCAAAATCGCATATATCTGCGCAAGGCAATTTATGACGTTG 7764

Qy 1261 TGAATTTCTTACTTCTACATCCGGGTGAATTTACATGGGCTTTTGGCCCTTGCAT 1320

Db 7763 TGAATCATGACCTTCTTCAACCCCGGTGAATTTGAACCTGGTTCCTACGTCCTGCAT 7704

Qy 1321 GTACAGCATCATTTTTCATTTGCTGGGGGATACCTTGTATCTCAACAGTACTGAGGTTG 1380

Db 7703 GCACACCAACACCTTCTCATTTACTCGGTATACCTCTATCTACACAGTTATAGCGTTG 7644

Qy 1381 ATGTGATGCGCTTGGGCTTGAATTTCAACATGTGTGACAGTATATGTCTCTTGGCT 1440

Db 7643 TTGTGATGTCCTGCTTGGCTTGAATTTCAACAAATTCAGTGTACCTTTCTTGGCTT 7584

Qy 1441 GATGCGCAATATCAAGGGA AAAAGCGGCTTGGGATATCAAGATGTCATGTCATGCA 1500

Db 7583 GATGCGCAATATCAAGGGA AAAAGCGGCTTGGGATATCAAGATGTCATGTCATGCA 7524

Qy 1501 CATTTACCAAGATCACTCGAATTTGATGCGCGATGTGACGCTTAAACGTTAGCCATTCCC 1560

Db 7523 CATTTATGAGATCACTTCTATCTCATGATAAAATGTGCAATTTAAACGTTAGCCATTCCC 7464

Qy 1561 AGCGCTCAGTTCAATTCATTCATAAGATTTAAACATGTGACGATTTGGAACTTGGGT 1620

Db 7463 GTTGCCCAATTCACATTTAACCCCGATTTAAACCTCGAAGATTTGAAACTTGGGT 7404

Qy 1621 CACTTGAATGATTTTACGTCACCGGATGATGATGATGATGATGATGATGATGATGAT 1680

Db 7403 GACTATGATGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7344

Qy 1681 TTGAGTCTAA 1690

Db 7343 CTCGTGTAA 7334

RESULT 14
 AF064791 1893 bp DNA linear BCT 31-DEC-2001
 LOCUS
 DEFINITION
 Pasteurella multocida thymidylate synthase (thyA) gene, complete cds.
 ACCESSION
 AF064791
 VERSION
 AF064791.1 GI:18000047
 KEYWORDS
 SOURCE
 ORGANISM
 Pasteurella multocida
 Pasteurella multocida
 Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Pasteurella.
 REFERENCE
 Moreno J.A., Bosch, M., Badiola, I., Llagostera, M. and Barbe, J. (1993) (1 to 1893)
 TITLE
 Direct Submission
 JOURNAL
 Submitted (12-MAY-1998) Departamento de Genética y Microbiología (Unidad de Microbiología), Universidad Autónoma de Barcelona, Edificio Cn, Bellaterra, Barcelona 08193, Spain
 FEATURES
 1..1893
 /organism="Pasteurella multocida"
 /mol_type="genomic DNA"

gene
/strain="PM25"
/specific host="bovine"
/db_xref="taxon:747"
826..1677
/gene="chvA"
826..1677
/gene="chvA"
/EC number="2.1.1.45"
/codon start=1
/transl_table=1
/product="thymidylate synthase"
/protein_id="AAL54880.1"
/db_xref="GI:18000048"
/translation="MKHYLELCORIVDEGVWYENERTKRCITVLDJLYDVANNOF
PLITTRKSYKAAIAEPLFQYIRGNADPRLGTWDAENENSAANPHKRGTD
DMGRVYGVGRAMRKNGERTVQDLKIVNLNLSI DMGLMTGPEEFLGRLPC
MHTHTPSLIGDLYLTSTVRSQSDVPLGHNFIQYFPLIAMAQTEKKAKAYHKII
NAHTYEDQJDLMKNVQLKREPFPLPOLHINPEIKTLEDLETVTWMDPKVTVGCHDA
IKRPSV"

ORIGIN

Query Match 26.6%; Score 775.2; DB 1; Length 1893;
Best Local Similarity 67.9%; Pred. No. 7.8e-229;
Matches 1119; Conservative 2; Mismatches 505; Indels 22; Gaps 2;

46 TGAACCCGATTTGTTTGCATGCGCCCTCTAGCGGTGCGTGTGATGCTGATTT
49 TGAACCCGATTTTGAATCGCTCTACGCTTACGCTGTATGCTTATATATCT
106 GGTGGGTTTCTTTTGTGATGCGTGGCCCAATGCGCGAGTCCGCGGAGTGG
109 GCTGGGTTTCTTTTGTGCGGCTTGGCGTGAACCTGTCTAACCAACGAGAGTGG
166 TTGACGCGTGAAGTCTGACTTGTATTCGCGGCTTTTAGTGTATGATCGG
169 CTGGAACAAGACCAAGTGATAGCTGTATTTAAAGGTTTATGGGGTTTATTAG
226 TGGCCGATTTGTTTGTATGCTTCTTCAATTCATTTCTTACAGATCTCGCTA
229 TGGCCGATTTGTTTGTATGCTTCTTCAATTCATTTCTTACAGATCTCGCTA
286 TTATATCAAGTGTGACTGCGGCGCATGCTCTTCCACGCGGCTTATGGGTGATGAC
289 TTATATCCGTTTGGGAAGCGGATATGCTTCCATGCGGCTTAAATGGTATTTT
346 CGCATGTTGTGATGCGGCTTAAACCAACGACCTCTTGTGTGCGGATTTTGT
349 AGCCATGTATATCAAGCTTAAATCAAAACGCGGTTTCTGCAACCGCGGATTTGT
406 TGGCCCTTATGATGCTTGTGCGGATGCGATCGGTAATCTTATGATATGTA
409 GGCACCGTTATTCCTTGGCTTAGTATGGACCGATCGGCACTTTATCAATGAGA
466 ACTTGGGGAAGATGAGATGCTGCTTGGCTTTTGTATTCCTTATGCTGCGCACT
469 ATTGAGGCGGCTGTACCGATGCGCATGCGGAGCTTATTCATCAGCGGATTTT
526 GCGCGCATGCTTCAAGCTTATGATTCGCTTAGAAGCGGCTTCTGCTTAT
529 ACCACGCGATCCCTCAAGCTTATGAGCGGATTTAGAGCATTTGTCCTTAT
586 TCTTATTTGTTTATGTTAACTGCTGCGTAGGACGATCCGATCTTTTATGAC
589 CCGAATTGTTATCAAAAACCTGACCAATCGGTGCAAGCTTGTATTTCTTACT
646 TGAATACGATACCTTCCGCTCTTGTGAAATCGTCCGTGAGCCAGATGCTCAATTGG
649 CGGTACCGATTTCCGCTTATGATGATTTTCCGTGAGCGATGAGCCCAATGAG
706 TCT---GTTGCTGCTTCAATTCAGGCGCAATCTCTCTTACTTATGATGATCAT
709 CTATATTTTGTCAACACATCTCTATGGAGAAATTCCTCAGACCAATATTTCTAT

763 CGGATTTTGAATGATGTTGTTCTTACAGCGCGGTTGTATCAAGCCGTGTAGCAGC 822
769 TGGTGGGTAATCATGTTAGTGGCATATCAAGT-----CGGGTA 809
823 AAATAGGTAATGATGATGAAACAGATTTATGATCTTGTGACCGCATGCTGATCAGG 882
810 AAAAAGTAATTTATGAAACATTTATGAACTTTGCAACGATGATCTTGTATGAGG 869
883 TGTGGGTTGAAATGAAAGCAAGCGGCAAGGTTTGTGACTGTGATTAATGCCGATT 942
870 TGTGGGTTGAAATGAAAGCAAGCGGCAAGGTTTGTGACTGTGATTAATGCCGATT 929
943 GACCTACGATGTGGCAACATGATTTCTCTAGAGACTTACACGGAAGTTTGGAA 1002
930 CACCTACGATGTGGCAACATGATTTCTCTAGAGACTTACACGGAAGTTTGGAA 989
1003 AGGTGCGGTGCGGAGTGTGCTGCTATATGCTGCTTACGATTAATGCGCGGATTTTG 1062
990 AGCGCAATGCAAGATTTTATGTTATTCGCTGCTATGCAATGCGGCAATTTCCG 1049
1063 CCAATTAGTACCAAACTGGGATGCTTATGCAATTTAAACCAAGCATGGCTACAA 1122
1050 CAAACTGCGCACCAAACTGGGATGCTTATGCAATTTAAACCAAGCATGGCTACAA 1109
1123 TCGTTACCGTAAGTGAAGATGACAGGAGCGGTATGCTTACGAGTGAAGCTTG 1182
1110 TCCGACCGCAAGGACCATGATATGAGACGCTTTATGGGATCAAGGGGCTCATG 1169
1183 GGTAAAGCTGATGCTGATATGATACCAAGTGAAGAAAGTTTGTATGATTTGACCG 1242
1170 GCGCAACCGAAGGTAACCGTCTGATCAATGCGCAAAATGTCATAATCTGGCAA 1229
1243 TGGCGTGAATGACGAGGTAATCTTAACTTCAATTCGCGGTAATTTCAATGAG 1302
1230 CGGATGATGACCGGTAATGATGATGATGATGATGATGATGATGATGATGATG 1289
1303 GTGTTGCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1362
1290 TTGCTACGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1349
1363 CAACATGATGACGCTTATGATGATGATGATGATGATGATGATGATGATGATG 1422
1350 CACGATGATGACGCTTATGATGATGATGATGATGATGATGATGATGATGATG 1409
1423 TTATGTTCTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1482
1410 GTTACGCTTCTTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 1469
1483 CAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1542
1470 CAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1529
1543 AAAACGTAAGCAATTCGAGGCTGATGATGATGATGATGATGATGATGATGATG 1602
1530 AAAACGTAAGCAATTCGAGGCTGATGATGATGATGATGATGATGATGATGATG 1589
1603 GGAATTTGAAACTTGGTCACTTGAATGATTTGATGATGATGATGATGATGATG 1662
1590 AGATTTGAAACTTGGTCACTTGAATGATTTTAAAGTCAAGGCTATCAATGACAGA 1649
1663 TCTATTCATTAACCGCTTTCAGTCAA 1690
1650 TGCCATTAATATCCTTCTCTGTGTA 1677

RESULT 15
LOCUS U32772 11545 bp DNA linear BCF 29-MAY-1998
DEFINITION Haemophilus influenzae Rd section 87 of 163 of the complete genome.
ACCESSION U32772 L42023
VERSION U32772.1 GI:1573918
KEYWORDS
SOURCE Haemophilus influenzae Rd Km20

PID:882719 percent identity: 35.32; identified by sequence similarity; putative"

/transl_table=1
/codon_start=1
/product="thymidylate synthetase (thyA)"
/protein_id="AAC22564.1"
/db_xref="GI:1573924"

/translation="MKOYLICRRIYVSEGWANERCKHCLTVINADLEVDVNNOF
PLITRKSYWKALAEFLGIRGIDNADFPALGKTITDANANEMANLPHRGSD
DMGRVYGVQGRAMREKNGEITIDLKRIVNNLTKGIDRKGILQITFNGEPLGLRC
MHTTPTSLVGDTHLITSYQRSCDPLGAFNQIQVFPFLAMQITKKGAKVHKIV
NAHIYEDOLELMDVQLKREBPPLPKLEINPDIKLIEDLETFMDDFKVGYOSHEP
IKYPSV"

4952 5473
/gene="H10906"
/gene="H10906"
4952 5473
/gene="H10906"
/note="similar to SP:P30134 PID:296183 PID:987637
PID:1031148 GB:U00096 percent identity: 57.41; identified
by sequence similarity; putative"

/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAC22565.1"
/db_xref="GI:1573925"

/translation="MDAAKVRSEDEKMMRYALIELADKRAALGRIPIVGAIVDARNI
IGEMNLSIVOSDPTAHAEIILRNGAKNIQNYRLNLTLYLEPCTMGALHSR
IKRLVAGSDYKTAIGSRFHFDDYKNNHLLEVTGVLAECSGKSTFQKREBK
KIERALLKSLSDK"

complement(5501, 5815)
/gene="H10907"
complement(5501, 5815)
/gene="H10907"
/note="hypothetical protein; identified by Genemark;
putative"

/codon_start=1
/transl_table=1
/product="H. influenzae predicted coding region H10907"
/protein_id="AAC22569.1"
/db_xref="GI:1573930"

/translation="MENKAERYOKAVNITDVLQSPFAKIIKKGLAINEINOKENRIF
POEFHFKRIGNMTDMSIFIEPANAIVRGILFRTLELLKLIQEEFPQVTFEITLNP
GR"

5895.6197
/gene="H10908"
5895.6197
/gene="H10908"
/note="hypothetical protein; identified by Genemark;

Query Match 25.6%; Score 744.2; DB 1; Length 11545;
Best Local Similarity 66.7%; Pred.No.4.9e-219;
Matches 1134; Conservative 0; Mismatches 538; Indels 27; Gaps 4;

QY 2 AGAAGGTTGTTATGCTCAGGGTATCTGACGTTTCCCATATTTGACCCCGTATTTGTT 61
DB 3273 ATAGGAAATTTATGAAATTCATTTACTTCTCCCACTTGGATCCGATATTTT 3332
QY 62 TCGATGGC-----CCTTAGCGGTCGCGTGTATGCTTATGTTGGTGGTTTC 115
DB 3333 ACCCTTGGCGATAGTATATCGCTTACGTTGATGAGCTTATGATACCTTTTAGGTTT 3392
QY 116 CTTTGGCTATGTTGGCAATCGCCGAGCGATCGCGCGGAGAGTGTGACGCGT 175
DB 3393 GTTTTGGACAGTGGCTTGGCTTGGCGCTAAATCGCCCAAAATAGCGGTTGGACAGT 3452
QY 176 GAGCAAGTCTGACTGTTGTTATGCGCGGCTTTTAAAGTGTAGTATCGTGGCGAGTT 235
DB 3453 GATCAAGTTGATGCTTACTTTCAACGGTTTATGGGGGTGTATTATGGCGGAGTGT 3512
QY 236 GGTATGATCTTACCAATTTGATCTGTCTGCTGACCGCTTATTTATTTATTTCAA 295
DB 3513 GCGGATGATTTTCTATATCTGATCATTTCTTCAAGAAACACTTTATTTATTCGCG 3572

QY 296 GTGTGACTGGCGGATGTCCTTTCACGCGGCTTATTTGGTGTATCATCCGCATGTTTC 355
DB 3573 GTTTGGAAGTGGATGATGCTTCCACGCTGGCTTATTTGGTATATGTTGCTATGAT 3632
QY 356 TGGTATCGCGTAACCAACGACCTCTTGGTGTGCGGATTTTGGCCCTTTA 415
DB 3633 TGGACATCTTATCTCAAAAGCTATTTTGGCAACGCGGATTTTGTGGCGCTTGG 3692
QY 416 GTGCCATTTGGTTGGGAATGGAGCATCGTAACTTATGAAATGAACTTTGGGA 475
DB 3693 ATTCCGTTGGTTTGGTTTATGGCAAAATGGTAAATTCATTAATCTTGAATAGGGA 3752
QY 476 CGAGTAACGATGTCGCTTGGGCTTTGATTTCCCTATGTTGGCC---CACTGCCGCG 532
DB 3753 CGGAAACGAATGTCGCTTGGCAATGATTTTCCGAATGATCTCTTTTACGCTCGT 3812
QY 533 CATCTTCACGCTTATGAAATTCGCTTAGAAGGCGTGTCTGTTTATCTTAT 592
DB 3813 CATCTTCACGCTTATGAAAGCTTTTGAAGGCTGTGTGTTTACGATTCGAT 3872
QY 593 TGGTTTATTTGTAACCTGTCGCTAGGACGATTCGGAATGTTTATGCTGATAC 652
DB 3873 ATTTTATTAATAAACACGTCATGCTCTGTTGACGTTTATTTCTTAATGCTAT 3932
QY 653 GGTACATTCGCTCTCTTGTGAATACGTCGTAAGCCAGATGCTCAGTTGGTGT 712
DB 3933 GGGGTCTCCGTTTATTTGTGAATATGCGGTAACCTGAAG-----TTGAAATTTTC 3986
QY 713 GTGCGCTCATTTCAATGGGCGCAATCTCTCTTACCTTATGATGATCATCGTAATTTG 772
DB 3987 TTTGGATATATACAGAGGCGAGCCCTTGTGTCGATGATATTTGTTGGTGTTC 4046
QY 773 ATGATGTTGTTGTTTCAAGCGGCTTGTATCAAGACCGTGTAGACCAAAATAGGT 832
DB 4047 ATTTATGCTTGGCTTATTCACGC-----AAAAGTCGCTATTAATAATGAG 4094
QY 833 AGTTAGTGAAACAGATTTAGATCTTGTACAGCGCATCGTATCAAGTGTGGT 892
DB 4095 ATTTTATGAGCAATATCTGAGCTTGTGCGCGATTTGTAAGTAAGGAGTGGT 4154
QY 893 GAAATGAAAGAAAGCGGCAAGCTTGTGATCTGTATTAATGCCATTTGACCTACAT 952
DB 4155 GCTAATGAAGTACAGTAAGCATTCCTCACGCTCATTAATGACATTTAGAAATGAT 4214
QY 953 GTGGGCAACATCAGTTCTCTAGTACATCAACGAGATTTTGAAGTGGCGTA 1012
DB 4215 GTGGCAATATCAATTTCCGCTGATTACTACCCGTAAGATTTTGAAGACGCGATT 4274
QY 1013 GCCGATTTGCTCGCTATATTTCTGTGTACGATTAATGCGCGGATTTGCCATTTAGT 1072
DB 4275 GCTGAATTTTATGTTATATCTGTGATATGACATGCGCTGATTTCCGCGACTTGGC 4334
QY 1073 ACCAAACCTGGATGCTAATGCCATTTAAACCAAGCATNGGCTCAACATCCTTACGCT 1132
DB 4335 ACGAAACCTGGATGCTAATGCCATTTAAACCAAGCATNGGCTCAACATCCTTACGCT 4394
QY 1133 AAAGTGAAGATGACATGAGACGCGTATGTTCAAGGTAAGACCTTGGGCTAAGCCT 1192
DB 4395 AGAGCGTTATGATATAGGGGCGGTATATGTTGTCAGAGGACATGCGCTAAGCCT 4454
QY 1193 GATGTGTGATATGACAGATTGAAAAGATTTGATGATTTGACCGCGCTGTAT 1252
DB 4455 AATGAGAAATCTATGATGATGCTACGTAAGTAAATGTTAAATACTTAAGAAAGTATGAT 4514
QY 1253 GACCGAGTGAAATTTCTTCAATCCGCGTGAATTTCAATGCGGCTTTTGGCG 1312
DB 4515 GATAGAGAGAGATTTTAACTTTTAAACCTCGGGGAATTTGATCTTGGTGTCTTGG 4574
QY 1313 CTTTGCATGTACAGCATTTTTCATTTGCTGGGGAATCCTTGTATCTCAACACTACT 1372
DB 4575 CTTTGTATGATACGATCTTTTCTCTGTGGGGAATCTTATCATCTTACTACTAT 4634
QY 1373 CAGGTTCACTGATGATGCGCTTGGGCTTGAATTTCAACATGTTGAGGTTATGTTCT 1432


```

Db 4635 CAGCGTTCCTGATGATGTCCTGCGCTTGATGAAATTCATCAATTCAGGTGTTACCTTC 4694
QY 1433 CTGCGCTGATGGACAGATCACAGGAAAAAGCCGGGCTTGCGGTATCACAGATCGTC 1492
Db 4695 TTAGCACTTATGGCTCAGATCACGGCAAAAAGCCGGGAAGCATATCATTAATTTGTG 4754
QY 1493 AATGCGCAATTTACCAAGATCACTCGAATTTGATGCGGATGTGCGAGCTAAACGTGAG 1552
Db 4755 AATGCGCAATTTATGAAGATCAGCTTGATGATGCGTGAAGTAAACGCGAG 4814
QY 1553 CCATTCCAGCGGCTCAGTTCATATCAATCCAAAGATTAAACACTGAGGATTTGGA 1612
Db 4815 CTTTCCCATTTACCAAAATTTAGAAATTAATCCAGATATAAAGCGCTTGAAGATTTGGA 4874
QY 1613 ACTTGGGTCACTTTGGATGATTTTGACGTACCGGATATCAGTTCCAGATCTTATCAA 1672
Db 4875 ACTTGGGTCAAGATGATTTTAAGTCGTTGGCTATCAATCCACGAACCAATTA 4934
QY 1673 TACCGTTTTCAGCTAAT 1691
Db 4935 TATCCTTTTCCGCTAAT 4953

```

Search completed: July 31, 2004, 15:38:59
 Job time : 7462 secs

